EXPRESS MAIL LABEL #EL 603865577 US

10/573229
IAP20 Rec'd PCT/PTO 24 MAR 2006

VERIFICATION OF TRANSLATION

Patent Application No. PCT/EP2004/010697
by Ganymed Pharmaceuticals AG

I, Dr. Georg Schnappauf
of c/o Dr. Volker Vossius
Patent and Law Firm
Geibelstrasse 6
81679 Munich
Germany

am the translator of the documents attached and I state that the following is a true translation to the best of my knowledge and belief of International Patent Application No. PCT/EP2004/010697 filed on September 23, 2004.

DATED: March 16, 2006

(Signature of translator)

Identification of Tumour-Associated Cell Surface Antigens for Diagnosis and Therapy

Despite interdisciplinary approaches and exhaustive use of classical therapeutic procedures, cancers are still among the leading causes of death. More recent therapeutic concepts aim at incorporating the patient's immune system into the overall therapeutic concept by using recombinant tumor vaccines and other specific measures such as antibody therapy. A prerequisite for the success of such a strategy is the recognition of tumor-specific or tumor-associated antigens or epitopes by the patient's immune system whose effector functions are to be interventionally enhanced. Tumor cells biologically differ substantially from their nonmalignant cells of origin. These differences are due to genetic alterations acquired during tumor development and result, inter alia, also in the formation of qualitatively or quantitatively altered molecular structures in the cancer cells. Tumor-associated structures of this kind which are recognized by the specific immune system of the tumor-harboring host are referred to as tumor-associated antigens. The specific recognition of tumor-associated antigens involves cellular and humoral mechanisms which are two functionally interconnected units: $CD4^+$ and $CD8^+$ T lymphocytes recognize the processed antigens presented on the molecules of the MHC (major histocompatibility complex) classes II and I, respectively, while B lymphocytes produce circulating antibody molecules which bind directly to unprocessed antigens. The potential clinical-therapeutical importance of tumor-associated antigens results from the fact that the recognition of antigens on neoplastic cells by the immune system leads to the initiation of cytotoxic effector mechanisms and, in the presence of T helper cells, can cause elimination of the cancer cells

(Pardoll, *Nat. Med.* 4:525-31, 1998). Accordingly, a central aim of tumor immunology is to molecularly define these structures. The molecular nature of these antigens has been enigmatic for a long time. Only after development of appropriate cloning techniques has it been possible to screen cDNA expression libraries of tumors systematically for tumor-associated antigens by analyzing the target structures of cytotoxic T lymphocytes (CTL) (van der Bruggen et al., *Science* 254:1643-7, 1991) or by using circulating autoantibodies (Sahin et al., *Curr. Opin. Immunol.* 9:709-16, 1997) as probes. To this end, cDNA expression libraries were prepared from fresh tumor tissue and recombinantly expressed as proteins in suitable systems. Immunoeffectors isolated from patients, namely CTL clones with tumor-specific lysis patterns, or circulating autoantibodies were utilized for cloning the respective antigens.

In recent years a multiplicity of antigens have been defined in various neoplasias by these approaches. The class of cancer/testis antigens (CTA) is of great interest here. CTA and genes encoding them (cancer/testis genes or CTG) are defined by their characteristic expression pattern [Tureci et al, *Mol Med Today*. 3:342-9, 1997]. They are not found in normal tissues, except testis and germ cells, but are expressed in a number of human malignomas, not tumor type-specifically but with different frequency in tumor entities of very different origins (Chen & Old, *Cancer J. Sci. Am.* 5:16-7, 1999). Serum reactivities against CTA are also not found in healthy controls but only in tumor patients. This class of antigens, in particular owing to its tissue distribution, is particularly valuable for immunotherapeutic projects and is tested in current clinical patient studies (Marchand et al., *Int. J. Cancer* 80:219-30, 1999; Knuth et al., *Cancer Chemother. Pharmacol.* 46:p46-51, 2000).

However, the probes utilized for antigen identification in the classical methods illustrated above are immunoeffectors (circulating autoantibodies or CTL clones) from patients usually having already advanced cancer. A number of data indicate that tumors can lead, for example, to tolerization and anergization of T cells and that, during the course of the disease, especially those specificities which could cause effective immune recognition are lost from the immunoeffector repertoire. Current patient studies have not yet produced any solid evidence of a real action of the previously found and utilized tumor-associated antigens. Accordingly, it cannot be ruled out that proteins evoking spontaneous immune responses are the wrong target structures.

It was the object of the present invention to provide target structures for a diagnosis and therapy of cancers.

According to the invention, this object is achieved by the subject matter of the claims.

According to the invention, a strategy for identifying and providing antigens expressed in association with a tumor and the nucleic acids coding therefor was pursued. This strategy is based on the evaluation of human protein and nucleic acid data bases with respect to potential cancer-specific antigens which are accessible on the cell surface. The definition of the filter criteria which are necessary for this together with a high throughput methodology for analysing all proteins, if possible, form the central part of the invention. Data mining first produces a list which is as complete as possible of all known genes which according to the basic principle "gene to mRNA to protein" are examined for the presence of one or more transmembrane domains. This is followed by a homology search, a classification of the hits in tissue specific

groups (among others tumor tissue) and an inspection of the real existence of the mRNA. Finally, the proteins which are identified in this manner are evaluated for their aberrant activation in tumors, e.g. by expression analyses and protein chemical procedures.

Data mining is a known method of identifying tumor-associated genes. In the conventional strategies, however, transcriptoms of normal tissue libraries are usually subtracted electronically from tumor tissue libraries, with the assumption that the remaining genes are tumor-specific (Schmitt et al., *Nucleic Acids Res.* 27:4251-60, 1999; Vasmatzis et al., *Proc. Natl. Acad. Sci. USA.* 95:300-4, 1998; Scheurle et al., *Cancer Res.* 60:4037-43, 2000).

The concept of the invention, however, is based on utilizing data mining for electronically extracting all genes coding for cancer specific antigens which are accessible on the cell surfaces and then evaluating said genes for ectopic expression in tumors.

The invention thus relates in one aspect to a strategy for identifying genes differentially expressed in tumors. Said strategy combines data mining of public sequence libraries ("*in silico*") with subsequent evaluating laboratory-experimental ("wet bench") studies.

According to the invention, a combined strategy based on different bioinformatic scripts enabled new genes coding for cancer specific antigens which are accessible on the cell surfaces to be identified. According to the invention, these tumor-associated genes and the genetic products encoded thereby were identified and provided independently of an immunogenic action.

The tumor-associated antigens identified according to

the invention have an amino acid sequence encoded by a nucleic acid which is selected from the group consisting of (a) a nucleic acid which comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45, 49, 53, 57, 61, 65, 69, 73, 77, 81, 85, 89, 93, 97, 101, 105, 109, 113, 117, 121, 125, 129, 133, 137, 141, 145, 149, 153, 157, 161, 165, 169, 173, 175, 179, 183, 187, 191, 195, 199, 203, 207, 211, 215, 219, 223, 227, 231, 235, 239, 243, 247, 251, 255, 259, 263, 267, 269, 271, 273, 275, 277, 279, 309 of the sequence listing, a part or derivative thereof, (b) a nucleic acid which hybridizes with the nucleic acid of (a) under stringent conditions, (c) a nucleic acid which is degenerate with respect to the nucleic acid of (a) or (b), and (d) a nucleic acid which is complementary to the nucleic acid of (a), (b) or (c). In a preferred embodiment, a tumor-associated antigen identified according to the invention has an amino acid sequence encoded by a nucleic acid which is selected from the group consisting of SEQ ID NOs: 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45, 49, 53, 57, 61, 65, 69, 73, 77, 81, 85, 89, 93, 97, 101, 105, 109, 113, 117, 121, 125, 129, 133, 137, 141, 145, 149, 153, 157, 161, 165, 169, 173, 175, 179, 183, 187, 191, 195, 199, 203, 207, 211, 215, 219, 223, 227, 231, 235, 239, 243, 247, 251, 255, 259, 263, 267, 269, 271, 273, 275, 277, 279, 309 of the sequence listing. In a further preferred embodiment, a tumor-associated antigen identified according to the invention comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 2,6, 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, 66, 70, 74, 78, 82, 86, 90, 94, 98, 102, 106, 110, 114, 118, 122, 126, 130, 134, 138, 142, 146, 150, 154, 158, 162, 166, 170, 174, 176, 180, 184, 188, 192, 196, 200, 204, 208, 212, 216, 220, 224, 228, 232, 236, 240, 244, 248, 252, 256, 260, 264, 268, 270, 272, 274, 276, 278, 280 to 308, 310 of the sequence listing, a part or derivative thereof.

The present invention generally relates to the use of tumor-associated antigens identified according to the invention or of parts thereof, of nucleic acids coding therefor or of nucleic acids directed against said coding nucleic acids or of antibodies directed against the tumor-associated antigens identified according to the invention or parts thereof for therapy and diagnosis. This utilization may relate to individual but also to combinations of two or more of these antigens, functional fragments, nucleic acids, antibodies, etc., in one embodiment also in combination with other tumor-associated genes and antigens for diagnosis, therapy and progress control.

The property of the tumor-associated antigens identified according to the invention that they are localized on or at the cell surface qualifies them as suitable targets or means for therapy and diagnosis. Especially suitable for this is a part of the tumor-associated antigens identified according to the invention which corresponds to the non-transmembrane portion, in particular the extracellular portion of the antigens, or is comprised thereof. Therefore, according to the invention, a part of the tumor-associated antigens identified according to the invention which corresponds to the non-transmembrane portion of the antigens or is comprised thereof, or a corresponding part of the nucleic acids coding for the tumor-associated antigens identified according to the invention is preferred for therapy or diagnosis. Similarly, the use of antibodies is preferred which are directed against a part of the tumor-associated antigens identified according to the invention which corresponds to the non-transmembrane portion of the antigens or is comprised thereof.

Preferred diseases for a therapy and/or diagnosis are those in which one or more of the tumor-associated

antigens identified according to the invention are selectively expressed or abnormally expressed.

The invention also relates to nucleic acids and genetic products which are expressed in association with a tumor cell and which are produced by altered splicing (splice variants) of genes or by altered translation with utilization of alternative open reading frames. Said nucleic acids comprise the sequences according to SEQ ID NOs: 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45, 49, 53, 57, 61, 65, 69, 73, 77, 81, 85, 89, 93, 97, 101, 105, 109, 113, 117, 121, 125, 129, 133, 137, 141, 145, 149, 153, 157, 161, 165, 169, 173, 175, 179, 183, 187, 191, 195, 199, 203, 207, 211, 215, 219, 223, 227, 231, 235, 239, 243, 247, 251, 255, 259, 263, 267, 269, 271, 273, 275, 277, 279, 309 of the sequence listing. Furthermore, the genetic products comprise all sequences according to SEQ ID NOs: 2,6, 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, 66, 70, 74, 78, 82, 86, 90, 94, 98, 102, 106, 110, 114, 118, 122, 126, 130, 134, 138, 142, 146, 150, 154, 158, 162, 166, 170, 174, 176, 180, 184, 188, 192, 196, 200, 204, 208, 212, 216, 220, 224, 228, 232, 236, 240, 244, 248, 252, 256, 260, 264, 268, 270, 272, 274, 276, 278, 280 to 308, 310 of the sequence listing. The splice variants of the invention can be used according to the invention as targets for diagnosis and therapy of tumor diseases.

Very different mechanisms may cause splice variants to be produced, for example

- utilization of variable transcription initiation sites
- utilization of additional exons
- complete or incomplete splicing out of single or two or more exons,
- splice regulator sequences altered via mutation (deletion or generation of new donor/acceptor sequences),
- incomplete elimination of intron sequences.

Altered splicing of a gene results in an altered transcript sequence (splice variant). Translation of a splice variant in the region of its altered sequence results in an altered protein which may be distinctly different in the structure and function from the original protein. Tumor-associated splice variants may produce tumor-associated transcripts and tumor-associated proteins/antigens. These may be utilized as molecular markers both for detecting tumor cells and for therapeutic targeting of tumors. Detection of tumor cells, for example in blood, serum, bone marrow, sputum, bronchial lavage, bodily secretions and tissue biopsies, may be carried out according to the invention, for example, after extraction of nucleic acids by PCR amplification with splice variant-specific oligonucleotides. According to the invention, all sequence-dependent detection systems are suitable for detection. These are, apart from PCR, for example gene chip/microarray systems, Northern blot, RNAse protection assays (RDA) and others. All detection systems have in common that detection is based on a specific hybridization with at least one splice variant-specific nucleic acid sequence. However, tumor cells may also be detected according to the invention by antibodies which recognize a specific epitope encoded by the splice variant. Said antibodies may be prepared by using for immunization peptides which are specific for said splice variant. Suitable for immunization are particularly the amino acids whose epitopes are distinctly different from the variant(s) of the genetic product, which is (are) preferably produced in healthy cells. Detection of the tumor cells with antibodies may be carried out here on a sample isolated from the patient or as imaging with intravenously administered antibodies.

In addition to diagnostic usability, splice variants having new or altered epitopes are attractive targets

for immunotherapy. The epitopes of the invention may be utilized for targeting therapeutically active monoclonal antibodies or T lymphocytes. In passive immunotherapy, antibodies or T lymphocytes which recognize splice variant-specific epitopes are adoptively transferred here. As in the case of other antigens, antibodies may be generated also by using standard technologies (immunization of animals, panning strategies for isolation of recombinant antibodies) with utilization of polypeptides which include these epitopes. Alternatively, it is possible to utilize for immunization nucleic acids coding for oligo- or polypeptides which contain said epitopes. Various techniques for in vitro or in vivo generation of epitope-specific T lymphocytes are known and have been described in detail (for example Kessler JH, et al. 2001, Sahin et al., 1997) and are likewise based on utilizing oligo- or polypeptides which contain the splice variant-specific epitopes or nucleic acids coding for said oligo- or polypeptides. Oligo- or polypeptides which contain the splice variant-specific epitopes or nucleic acids coding for said polypeptides may also be used for utilization as pharmaceutically active substances in active immunotherapy (vaccination, vaccine therapy).

The aberrant expression of genes in tumor cells also can be due to an altered methylation pattern of their promoters (De Smet C et al., Mol. Cell Biol. 24(11):4781-90, 2004; De Smet C et al., Mol. Cell Biol. 19(11):7327-35, 1999; De Smet C et al., Proc. Natl. Acad. Sci. U S A. 93(14):7149-53, 1996). These differences in methylation can be used as indirect markers for the condition of the respective gene changed in the tumor. Accordingly, the increase or decrease of base methylations within the promoter region can be used for diagnostic purposes.

In one aspect, the invention relates to a

pharmaceutical composition comprising an agent which recognizes the tumor-associated antigen identified according to the invention and which is preferably selective for cells which have expression or abnormal expression of a tumor-associated antigen identified according to the invention. In particular embodiments, said agent may cause induction of cell death, reduction in cell growth, damage to the cell membrane or secretion of cytokines and preferably have a tumor-inhibiting activity. In one embodiment, the agent is an antisense nucleic acid which hybridizes selectively with the nucleic acid coding for the tumor-associated antigen. In a further embodiment, the agent is an antibody which binds selectively to the tumor-associated antigen, in particular a complement-activated antibody which binds selectively to the tumor-associated antigen. In a further embodiment, the agent comprises two or more agents which each selectively recognize different tumor-associated antigens, at least one of which is a tumor-associated antigen identified according to the invention. Recognition needs not be accompanied directly with inhibition of activity or expression of the antigen. In this aspect of the invention, the antigen selectively limited to tumors preferably serves as a label for recruiting effector mechanisms to this specific location. In a preferred embodiment, the agent is a cytotoxic T lymphocyte which recognizes the antigen on an HLA molecule and lyses the cell labeled in this way. In a further embodiment, the agent is an antibody which binds selectively to the tumor-associated antigen and thus recruits natural or artificial effector mechanisms to said cell. In a further embodiment, the agent is a T helper lymphocyte which enhances effector functions of other cells specifically recognizing said antigen.

In one aspect, the invention relates to a pharmaceutical composition comprising an agent which inhibits expression or activity of a tumor-associated

antigen identified according to the invention. In a preferred embodiment, the agent is an antisense nucleic acid which hybridizes selectively with the nucleic acid coding for the tumor-associated antigen. In a further embodiment, the agent is an antibody which binds selectively to the tumor-associated antigen. In a further embodiment, the agent comprises two or more agents which each selectively inhibit expression or activity of different tumor-associated antigens, at least one of which is a tumor-associated antigen identified according to the invention.

The activity of a tumor-associated antigen identified according to the invention can be any activity of a protein or a peptide. Thus, the therapeutic and diagnostic methods according to the invention can also aim at inhibiting or reducing this activity or testing this activity.

The invention furthermore relates to a pharmaceutical composition which comprises an agent which, when administered, selectively increases the amount of complexes between an HLA molecule and a peptide epitope from the tumor-associated antigen identified according to the invention. In one embodiment, the agent comprises one or more components selected from the group consisting of (i) the tumor-associated antigen or a part thereof, (ii) a nucleic acid which codes for said tumor-associated antigen or a part thereof, (iii) a host cell which expresses said tumor-associated antigen or a part thereof, and (iv) isolated complexes between peptide epitopes from said tumor-associated antigen and an MHC molecule. In one embodiment, the agent comprises two or more agents which each selectively increase the amount of complexes between MHC molecules and peptide epitopes of different tumor-associated antigens, at least one of which is a tumor-associated antigen identified according to the invention.

The invention furthermore relates to a pharmaceutical composition which comprises one or more components selected from the group consisting of (i) a tumor-associated antigen identified according to the invention or a part thereof, (ii) a nucleic acid which codes for a tumor-associated antigen identified according to the invention or for a part thereof, (iii) an antibody which binds to a tumor-associated antigen identified according to the invention or to a part thereof, (iv) an antisense nucleic acid which hybridizes specifically with a nucleic acid coding for a tumor-associated antigen identified according to the invention, (v) a host cell which expresses a tumor-associated antigen identified according to the invention or a part thereof, and (vi) isolated complexes between a tumor-associated antigen identified according to the invention or a part thereof and an HLA molecule.

A nucleic acid coding for a tumor-associated antigen identified according to the invention or for a part thereof may be present in the pharmaceutical composition in an expression vector and functionally linked to a promoter.

A host cell present in a pharmaceutical composition of the invention may secrete the tumor-associated antigen or the part thereof, express it on the surface or may additionally express an HLA molecule which binds to said tumor-associated antigen or said part thereof. In one embodiment, the host cell expresses the HLA molecule endogenously. In a further embodiment, the host cell expresses the HLA molecule and/or the tumor-associated antigen or the part thereof in a recombinant manner. The host cell is preferably nonproliferative. In a preferred embodiment, the host cell is an antigen-presenting cell, in particular a dendritic cell, a monocyte or a macrophage.

An antibody present in a pharmaceutical composition of the invention may be a monoclonal antibody. In further embodiments, the antibody is a chimeric or humanized antibody, a fragment of a natural antibody or a synthetic antibody, all of which may be produced by combinatory techniques. The antibody may be coupled to a therapeutically or diagnostically useful agent.

An antisense nucleic acid present in a pharmaceutical composition of the invention may comprise a sequence of 6-50, in particular 10-30, 15-30 and 20-30, contiguous nucleotides of the nucleic acid coding for the tumor-associated antigen identified according to the invention.

In further embodiments, a tumor-associated antigen, provided by a pharmaceutical composition of the invention either directly or via expression of a nucleic acid, or a part thereof binds to MHC molecules on the surface of cells, said binding preferably causing a cytolytic response and/or inducing cytokine release.

A pharmaceutical composition of the invention may comprise a pharmaceutically compatible carrier and/or an adjuvant. The adjuvant may be selected from saponin, GM-CSF, CpG oligonucleotides, RNA, a cytokine or a chemokine. A pharmaceutical composition of the invention is preferably used for the treatment of a disease characterized by selective expression or abnormal expression of a tumor-associated antigen. In a preferred embodiment, the disease is cancer.

The invention furthermore relates to methods of treating, diagnosing or monitoring, i.e. determining the regression, progression and/or onset of, a disease characterized by expression or abnormal expression of one of more tumor-associated antigens.

In one embodiment, the methods of treatment according to the invention comprise administering a pharmaceutical composition of the invention.

The methods of diagnosing and/or methods of monitoring according to the invention generally concern the use of means for the detection and/or the determination and/or the monitoring of the quantity of (i) a nucleic acid, which codes for the tumor-associated antigen, or a part thereof and/or (ii) the tumor-associated antigen or a part thereof and/or (iii) an antibody against the tumor-associated antigen or a part thereof and/or (iv) cytotoxic or T helper lymphocytes, which are specific for the tumor-associated antigen or a part thereof, in a biologic sample isolated from a patient.

In one aspect, the invention relates to a method of diagnosing a disease characterized by expression or abnormal expression of a tumor-associated antigen identified according to the invention. The method comprises (i) detection of a nucleic acid which codes for the tumor-associated antigen or of a part thereof and/or (ii) detection of the tumor-associated antigen or of a part thereof, and/or (iii) detection of an antibody to the tumor-associated antigen or to a part thereof and/or (iv) detection of cytotoxic or T helper lymphocytes which are specific for the tumor-associated antigen or for a part thereof in a biological sample isolated from a patient. In particular embodiments, detection comprises (i) contacting the biological sample with an agent which binds specifically to the nucleic acid coding for the tumor-associated antigen or to the part thereof, to said tumor-associated antigen or said part thereof, to the antibody or to cytotoxic or T helper lymphocytes specific for the tumor-associated antigen or parts thereof, and (ii) detecting the formation of a complex between the agent and the nucleic acid or the part thereof, the tumor-associated

antigen or the part thereof, the antibody or the cytotoxic or T helper lymphocytes. In one embodiment, the disease is characterized by expression or abnormal expression of two or more different tumor-associated antigens and detection comprises detection of two or more nucleic acids coding for said two or more different tumor-associated antigens or of parts thereof, detection of two or more different tumor-associated antigens or of parts thereof, detection of two or more antibodies binding to said two or more different tumor-associated antigens or to parts thereof or detection of two or more cytotoxic or T helper lymphocytes specific for said two or more different tumor-associated antigens. In a further embodiment, the biological sample isolated from the patient is compared to a comparable normal biological sample.

The methods of diagnosing according to the invention may also utilize altered methylation patterns of the promoter region of the respective tumor-associated gene product. The detection of such methylation patterns can be performed by using methods on the basis of PCR, with the aid of restriction enzymes or by sequencing. A test suitable for this can be as follows: (1) extraction of DNA from tissue samples of patients, for example using paraffin embedded material, (2) treatment of the DNA with bisulfite containing reagents (i.e. as described in Clark S.J. et al., Nucleic Acids Res. 22(15):2990-7, 1994), (3) amplification of DNA by means of PCR and (4) analysis by determining the amount of sequence specific amplification products (e.g. by means of quantitative PCR, hybridization techniques such as microarray methods).

The methods of diagnosing according to the invention can concern also the use of the tumor-associated antigens identified according to the invention as prognostic markers, in order to predict metastasis, e.g. through testing the migration behavior of cells,

and therefore a worsened course of the disease, whereby among other things planning of a more aggressive therapy is made possible.

In a further aspect, the invention relates to a method for determining regression, course or onset of a disease characterized by expression or abnormal expression of a tumor-associated antigen identified according to the invention, which method comprises monitoring a sample from a patient who has said disease or is suspected of falling ill with said disease, with respect to one or more parameters selected from the group consisting of (i) the amount of nucleic acid which codes for the tumor-associated antigen or of a part thereof, (ii) the amount of the tumor-associated antigen or a part thereof, (iii) the amount of antibodies which bind to the tumor-associated antigen or to a part thereof, and (iv) the amount of cytolytic T cells or T helper cells which are specific for a complex between the tumor-associated antigen or a part thereof and an MHC molecule. The method preferably comprises determining the parameter(s) in a first sample at a first point in time and in a further sample at a second point in time and in which the course of the disease is determined by comparing the two samples. In particular embodiments, the disease is characterized by expression or abnormal expression of two or more different tumor-associated antigens and monitoring comprises monitoring (i) the amount of two or more nucleic acids which code for said two or more different tumor-associated antigens or of parts thereof, and/or (ii) the amount of said two or more different tumor-associated antigens or of parts thereof, and/or (iii) the amount of two or more antibodies which bind to said two or more different tumor-associated antigens or to parts thereof, and/or (iv) the amount of two or more cytolytic T cells or of T helper cells which are specific for complexes between said two or more different tumor-associated antigens or of parts thereof

and MHC molecules.

According to the invention, detection of a nucleic acid or of a part thereof or determining or monitoring the amount of a nucleic acid or of a part thereof may be carried out using a polynucleotide probe which hybridizes specifically to said nucleic acid or said part thereof or may be carried out by selective amplification of said nucleic acid or said part thereof. In one embodiment, the polynucleotide probe comprises a sequence of 6-50, in particular 10-30, 15-30 and 20-30, contiguous nucleotides of said nucleic acid.

In certain embodiments of the methods of diagnosing of the invention, the promoter region or part thereof of a nucleic acid coding for a tumor-associated antigen identified according to the invention and being present in the form of genomic DNA is selectively amplified following treatment with a bisulfite containing reagent. The nucleic acid is preferably isolated from a sample of a patient to be examined before treatment with the bisulfite containing reagent. The oligonucleotides used in such amplification preferably have a sequence binding to the nucleic acid treated with a bisulfite containing reagent and preferably are completely complementary thereto. Preferably, the oligonucleotides are adapted to a different degree of methylation of the nucleic acid and bring about amplification products which can be differentiated.

According to the invention, detection of a tumor-associated antigen or of a part thereof or determining or monitoring the amount of a tumor-associated antigen or of a part thereof may be carried out using an antibody binding specifically to said tumor-associated antigen or said part thereof.

In certain embodiments, the tumor-associated antigen to

be detected or the part thereof is present in a complex with an MHC molecule, in particular an HLA molecule.

According to the invention, detection of an antibody or determining or monitoring the amount of antibodies may be carried out using a protein or peptide binding specifically to said antibody.

According to the invention, detection of cytolytic T cells or of T helper cells or determining or monitoring the amount of cytolytic T cells or of T helper cells which are specific for complexes between an antigen or a part thereof and MHC molecules may be carried out using a cell presenting the complex between said antigen or said part thereof and an MHC molecule.

The polynucleotide probe, the antibody, the protein or peptide or the cell, which is used for detection or determining or monitoring, is preferably labeled in a detectable manner. In particular embodiments, the detectable marker is a radioactive marker or an enzymic marker. T lymphocytes may additionally be detected by detecting their proliferation, their cytokine production, and their cytotoxic activity triggered by specific stimulation with the complex of MHC and tumor-associated antigen or parts thereof. T lymphocytes may also be detected via a recombinant MHC molecule or else a complex of two or more MHC molecules which are loaded with the particular immunogenic fragment of one or more of the tumor-associated antigens and which can identify the specific T lymphocytes by contacting the specific T cell receptor.

In a further aspect, the invention relates to a method of treating, diagnosing or monitoring a disease characterized by expression or abnormal expression of a tumor-associated antigen identified according to the invention, which method comprises administering an antibody which binds to said tumor-associated antigen

or to a part thereof and which is coupled to a therapeutic or diagnostic agent. The antibody may be a monoclonal antibody. In further embodiments, the antibody is a chimeric or humanized antibody or a fragment of a natural antibody.

The invention also relates to a method of treating a patient having a disease characterized by expression or abnormal expression of a tumor-associated antigen identified according to the invention, which method comprises (i) removing a sample containing immunoreactive cells from said patient, (ii) contacting said sample with a host cell expressing said tumor-associated antigen or a part thereof, under conditions which favor production of cytolytic T cells against said tumor-associated antigen or a part thereof, and (iii) introducing the cytolytic T cells into the patient in an amount suitable for lysing cells expressing the tumor-associated antigen or a part thereof. The invention likewise relates to cloning the T cell receptor of cytolytic T cells against the tumor-associated antigen. Said receptor may be transferred to other T cells which thus receive the desired specificity and, as under (iii), may be introduced into the patient.

In one embodiment, the host cell endogenously expresses an HLA molecule. In a further embodiment, the host cell recombinantly expresses an HLA molecule and/or the tumor-associated antigen or the part thereof. The host cell is preferably nonproliferative. In a preferred embodiment, the host cell is an antigen-presenting cell, in particular a dendritic cell, a monocyte or a macrophage.

In a further aspect, the invention relates to a method of treating a patient having a disease characterized by expression or abnormal expression of a tumor-associated antigen, which method comprises (i) identifying a

nucleic acid which codes for a tumor-associated antigen identified according to the invention and which is expressed by cells associated with said disease, (ii) transfecting a host cell with said nucleic acid or a part thereof, (iii) culturing the transfected host cell for expression of said nucleic acid (this is not obligatory when a high rate of transfection is obtained), and (iv) introducing the host cells or an extract thereof into the patient in an amount suitable for increasing the immune response to the patient's cells associated with the disease. The method may further comprise identifying an MHC molecule presenting the tumor-associated antigen or a part thereof, with the host cell expressing the identified MHC molecule and presenting said tumor-associated antigen or a part thereof. The immune response may comprise a B cell response or a T cell response. Furthermore, a T cell response may comprise production of cytolytic T cells and/or T helper cells which are specific for the host cells presenting the tumor-associated antigen or a part thereof or specific for cells of the patient which express said tumor-associated antigen or a part thereof.

The invention also relates to a method of treating a disease characterized by expression or abnormal expression of a tumor-associated antigen identified according to the invention, which method comprises (i) identifying cells from the patient which express abnormal amounts of the tumor-associated antigen, (ii) isolating a sample of said cells, (iii) culturing said cells, and (iv) introducing said cells into the patient in an amount suitable for triggering an immune response to the cells.

Preferably, the host cells used according to the invention are nonproliferative or are rendered nonproliferative. A disease characterized by expression or abnormal expression of a tumor-associated antigen is

in particular cancer.

The present invention furthermore relates to a nucleic acid selected from the group consisting of (a) a nucleic acid which comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45, 49, 53, 57, 61, 65, 69, 73, 77, 81, 85, 89, 93, 97, 101, 105, 109, 113, 117, 121, 125, 129, 133, 137, 141, 145, 149, 153, 157, 161, 165, 169, 173, 175, 179, 183, 187, 191, 195, 199, 203, 207, 211, 215, 219, 223, 227, 231, 235, 239, 243, 247, 251, 255, 259, 263, 267, 269, 271, 273, 275, 277, 279, 309 of the sequence listing, a part or derivative thereof, (b) a nucleic acid which hybridizes with the nucleic acid of (a) under stringent conditions, (c) a nucleic acid which is degenerate with respect to the nucleic acid of (a) or (b), and (d) a nucleic acid which is complementary to the nucleic acid of (a), (b) or (c). The invention furthermore relates to a nucleic acid, which codes for a protein or polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 2,6, 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, 66, 70, 74, 78, 82, 86, 90, 94, 98, 102, 106, 110, 114, 118, 122, 126, 130, 134, 138, 142, 146, 150, 154, 158, 162, 166, 170, 174, 176, 180, 184, 188, 192, 196, 200, 204, 208, 212, 216, 220, 224, 228, 232, 236, 240, 244, 248, 252, 256, 260, 264, 268, 270, 272, 274, 276, 278, 280 to 308, 310 of the sequence listing, a part or derivative thereof.

In a further aspect, the invention relates to promoter sequences of nucleic acids of the invention. These sequences may be functionally linked to another gene, preferably in an expression vector, and thus ensure selective expression of said gene in appropriate cells.

In a further aspect, the invention relates to a recombinant nucleic acid molecule, in particular DNA or RNA molecule, which comprises a nucleic acid of the

invention.

The invention also relates to host cells which contain a nucleic acid of the invention or a recombinant nucleic acid molecule comprising a nucleic acid of the invention.

The host cell may also comprise a nucleic acid coding for a HLA molecule. In one embodiment, the host cell endogenously expresses the HLA molecule. In a further embodiment, the host cell recombinantly expresses the HLA molecule and/or the nucleic acid of the invention or a part thereof. Preferably, the host cell is nonproliferative. In a preferred embodiment, the host cell is an antigen-presenting cell, in particular a dendritic cell, a monocyte or a macrophage.

In a further embodiment, the invention relates to oligonucleotides which hybridize with a nucleic acid identified according to the invention and which may be used as genetic probes or as "antisense" molecules. Nucleic acid molecules in the form of oligonucleotide primers or competent samples, which hybridize with a nucleic acid identified according to the invention or parts thereof, may be used for finding nucleic acids which are homologous to said nucleic acid identified according to the invention. PCR amplification, Southern and Northern hybridization may be employed for finding homologous nucleic acids. Hybridization may be carried out under low stringency, more preferably under medium stringency and most preferably under high stringency conditions. The term "stringent conditions" according to the invention refers to conditions which allow specific hybridization between polynucleotides.

In a further aspect, the invention relates to a protein or polypeptide which is encoded by a nucleic acid selected from the group consisting of (a) a nucleic acid which comprises a nucleic acid sequence selected

from the group consisting of SEQ ID NOs: 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45, 49, 53, 57, 61, 65, 69, 73, 77, 81, 85, 89, 93, 97, 101, 105, 109, 113, 117, 121, 125, 129, 133, 137, 141, 145, 149, 153, 157, 161, 165, 169, 173, 175, 179, 183, 187, 191, 195, 199, 203, 207, 211, 215, 219, 223, 227, 231, 235, 239, 243, 247, 251, 255, 259, 263, 267, 269, 271, 273, 275, 277, 279, 309 of the sequence listing, a part or derivative thereof, (b) a nucleic acid which hybridizes with the nucleic acid of (a) under stringent conditions, (c) a nucleic acid which is degenerate with respect to the nucleic acid of (a) or (b), and (d) a nucleic acid which is complementary to the nucleic acid of (a), (b) or (c). In a preferred embodiment, the invention relates to a protein or polypeptide which comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 2,6, 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, 66, 70, 74, 78, 82, 86, 90, 94, 98, 102, 106, 110, 114, 118, 122, 126, 130, 134, 138, 142, 146, 150, 154, 158, 162, 166, 170, 174, 176, 180, 184, 188, 192, 196, 200, 204, 208, 212, 216, 220, 224, 228, 232, 236, 240, 244, 248, 252, 256, 260, 264, 268, 270, 272, 274, 276, 278, 280 to 308, 310 of the sequence listing, a part or derivative thereof.

In a further aspect, the invention relates to an immunogenic fragment of a tumor-associated antigen identified according to the invention. Said fragment preferably binds to a human HLA receptor or to a human antibody. A fragment of the invention preferably comprises a sequence of at least 6, in particular at least 8, at least 10, at least 12, at least 15, at least 20, at least 30 or at least 50, amino acids.

In a further aspect, the invention relates to an agent which binds to a tumor-associated antigen identified according to the invention or to a part thereof. In a preferred embodiment, the agent is an antibody. In further embodiments, the antibody is a chimeric, a

humanized antibody or an antibody produced by combinatory techniques or is a fragment of an antibody. Furthermore, the invention relates to an antibody which binds selectively to a complex of (i) a tumor-associated antigen identified according to the invention or a part thereof and (ii) an MHC molecule to which said tumor-associated antigen identified according to the invention or said part thereof binds, with said antibody not binding to (i) or (ii) alone. An antibody of the invention may be a monoclonal antibody. In further embodiments, the antibody is a chimeric or humanized antibody or a fragment of a natural antibody.

The invention furthermore relates to a conjugate between an agent of the invention which binds to a tumor-associated antigen identified according to the invention or to a part thereof or an antibody of the invention and a therapeutic or diagnostic agent. In one embodiment, the therapeutic or diagnostic agent is a toxin.

In a further aspect, the invention relates to a kit for detecting expression or abnormal expression of a tumor-associated antigen identified according to the invention, which kit comprises agents for detection (i) of the nucleic acid which codes for the tumor-associated antigen or of a part thereof, (ii) of the tumor-associated antigen or of a part thereof, (iii) of antibodies which bind to the tumor-associated antigen or to a part thereof, and/or (iv) of T cells which are specific for a complex between the tumor-associated antigen or a part thereof and an MHC molecule. In one embodiment, the agents for detection of the nucleic acid or the part thereof are nucleic acid molecules for selective amplification of said nucleic acid, which comprise, in particular a sequence of 6-50, in particular 10-30, 15-30 and 20-30, contiguous nucleotides of said nucleic acid.

Detailed description of the invention

According to the invention, genes are described which are expressed in tumor cells selectively or aberrantly and which are tumor-associated antigens.

According to the invention, these genes or their derivatives are preferred target structures for therapeutic approaches. Conceptionally, said therapeutic approaches may aim at inhibiting the activity of the selectively expressed tumor-associated genetic product. This is useful, if said aberrant respective selective expression is functionally important in tumor pathogenecity and if its ligation is accompanied by selective damage of the corresponding cells. Other therapeutic concepts contemplate tumor-associated antigens as labels which recruit effector mechanisms having cell-damaging potential selectively to tumor cells. Here, the function of the target molecule itself and its role in tumor development are totally irrelevant.

"Derivative" of a nucleic acid means according to the invention that single or multiple nucleotide substitutions, deletions and/or additions are present in said nucleic acid. Furthermore, the term "derivative" also comprises chemical derivatization of a nucleic acid on a base, on a sugar or on a phosphate of a nucleotide. The term "derivative" also comprises nucleic acids which contain nucleotides and nucleotide analogs not occurring naturally.

According to the invention, a nucleic acid is preferably deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). Nucleic acids comprise according to the invention genomic DNA, cDNA, mRNA, recombinantly produced and chemically synthesized molecules. According to the invention, a nucleic acid may be present as a single-stranded or double-stranded and

linear or covalently circularly closed molecule.

The nucleic acids described according to the invention have preferably been isolated. The term "isolated nucleic acid" means according to the invention that the nucleic acid was (i) amplified *in vitro*, for example by polymerase chain reaction (PCR), (ii) recombinantly produced by cloning, (iii) purified, for example by cleavage and gel-electrophoretic fractionation, or (iv) synthesized, for example by chemical synthesis. An isolated nucleic acid is a nucleic acid which is available for manipulation by recombinant DNA techniques.

A nucleic acid is "complementary" to another nucleic acid if the two sequences are capable of hybridizing and forming a stable duplex with one another, with hybridization preferably being carried out under conditions which allow specific hybridization between polynucleotides (stringent conditions). Stringent conditions are described, for example, in Molecular Cloning: A Laboratory Manual, J. Sambrook et al., Editors, 2nd Edition, Cold Spring Harbor Laboratory press, Cold Spring Harbor, New York, 1989 or Current Protocols in Molecular Biology, F.M. Ausubel et al., Editors, John Wiley & Sons, Inc., New York and refer, for example, to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin, 2.5 mM NaH_2PO_4 (pH 7), 0.5% SDS, 2 mM EDTA). SSC is 0.15 M sodium chloride/0.15 M sodium citrate, pH 7. After hybridization, the membrane to which the DNA has been transferred is washed, for example, in 2 × SSC at room temperature and then in 0.1-0.5 × SSC/0.1 × SDS at temperatures of up to 68°C.

According to the invention, complementary nucleic acids have at least 40%, in particular at least 50%, at least 60%, at least 70%, at least 80%, at least 90% and

preferably at least 95%, at least 98% or at least 99%, identical nucleotides.

Nucleic acids coding for tumor-associated antigens may, according to the invention, be present alone or in combination with other nucleic acids, in particular heterologous nucleic acids. In preferred embodiments, a nucleic acid is functionally linked to expression control sequences or regulatory sequences which may be homologous or heterologous with respect to said nucleic acid. A coding sequence and a regulatory sequence are "functionally" linked to one another, if they are covalently linked to one another in such a way that expression or transcription of said coding sequence is under the control or under the influence of said regulatory sequence. If the coding sequence is to be translated into a functional protein, then, with a regulatory sequence functionally linked to said coding sequence, induction of said regulatory sequence results in transcription of said coding sequence, without causing a frame shift in the coding sequence or said coding sequence not being capable of being translated into the desired protein or peptide.

The term "expression control sequence" or "regulatory sequence" comprises according to the invention promoters, enhancers and other control elements which regulate expression of a gene. In particular embodiments of the invention, the expression control sequences can be regulated. The exact structure of regulatory sequences may vary as a function of the species or cell type, but generally comprises 5'untranscribed and 5'untranslated sequences which are involved in initiation of transcription and translation, respectively, such as TATA box, capping sequence, CAAT sequence, and the like. More specifically, 5'untranscribed regulatory sequences comprise a promoter region which includes a promoter sequence for transcriptional control of the

functionally linked gene. Regulatory sequences may also comprise enhancer sequences or upstream activator sequences.

Thus, on the one hand, the tumor-associated antigens illustrated herein may be combined with any expression control sequences and promoters. On the other hand, however, the promoters of the tumor-associated genetic products illustrated herein may, according to the invention, be combined with any other genes. This allows the selective activity of these promoters to be utilized.

According to the invention, a nucleic acid may furthermore be present in combination with another nucleic acid which codes for a polypeptide controlling secretion of the protein or polypeptide encoded by said nucleic acid from a host cell. According to the invention, a nucleic acid may also be present in combination with another nucleic acid which codes for a polypeptide causing the encoded protein or polypeptide to be anchored on the cell membrane of the host cell or compartmentalized into particular organelles of said cell.

In a preferred embodiment, a recombinant DNA molecule is according to the invention a vector, where appropriate with a promoter, which controls expression of a nucleic acid, for example a nucleic acid coding for a tumor-associated antigen of the invention. The term "vector" is used here in its most general meaning and comprises any intermediary vehicle for a nucleic acid which enables said nucleic acid, for example, to be introduced into prokaryotic and/or eukaryotic cells and, where appropriate, to be integrated into a genome. Vectors of this kind are preferably replicated and/or expressed in the cells. An intermediary vehicle may be adapted, for example, to the use in electroporation, in bombardment with microprojectiles, in liposomal

administration, in the transfer with the aid of agrobacteria or in insertion via DNA or RNA viruses. Vectors comprise plasmids, phagemids, bacteriophages or viral genomes.

The nucleic acids coding for a tumor-associated antigen identified according to the invention may be used for transfection of host cells. Nucleic acids here mean both recombinant DNA and RNA. Recombinant RNA may be prepared by in-vitro transcription of a DNA template. Furthermore, it may be modified by stabilizing sequences, capping and polyadenylation prior to application. According to the invention, the term "host cell" relates to any cell which can be transformed or transfected with an exogenous nucleic acid. The term "host cells" comprises according to the invention prokaryotic (e.g. *E. coli*) or eukaryotic cells (e.g. dendritic cells, B cells, CHO cells, COS cells, K562 cells, yeast cells and insect cells). Particular preference is given to mammalian cells such as cells from humans, mice, hamsters, pigs, goats, primates. The cells may be derived from a multiplicity of tissue types and comprise primary cells and cell lines. Specific examples comprise keratinocytes, peripheral blood leukocytes, stem cells of the bone marrow and embryonic stem cells. In further embodiments, the host cell is an antigen-presenting cell, in particular a dendritic cell, monocyte or a macrophage. A nucleic acid may be present in the host cell in the form of a single copy or of two or more copies and, in one embodiment, is expressed in the host cell.

According to the invention, the term "expression" is used in its most general meaning and comprises the production of RNA or of RNA and protein. It also comprises partial expression of nucleic acids. Furthermore, expression may be carried out transiently or stably. Preferred expression systems in mammalian cells comprise pcDNA3.1 and pRc/CMV (Invitrogen,

Carlsbad, CA), which contain a selective marker such as a gene imparting resistance to G418 (and thus enabling stably transfected cell lines to be selected) and the enhancer-promoter sequences of cytomegalovirus (CMV).

In those cases of the invention in which an HLA molecule presents a tumor-associated antigen or a part thereof, an expression vector may also comprise a nucleic acid sequence coding for said HLA molecule. The nucleic acid sequence coding for the HLA molecule may be present on the same expression vector as the nucleic acid coding for the tumor-associated antigen or the part thereof, or both nucleic acids may be present on different expression vectors. In the latter case, the two expression vectors may be cotransfected into a cell. If a host cell expresses neither the tumor-associated antigen or the part thereof nor the HLA molecule, both nucleic acids coding therefor are transfected into the cell either on the same expression vector or on different expression vectors. If the cell already expresses the HLA molecule, only the nucleic acid sequence coding for the tumor-associated antigen or the part thereof can be transfected into the cell.

The invention also comprises kits for amplification of a nucleic acid coding for a tumor-associated antigen. Such kits comprise, for example, a pair of amplification primers which hybridize to the nucleic acid coding for the tumor-associated antigen. The primers preferably comprise a sequence of 6-50, in particular 10-30, 15-30 and 20-30 contiguous nucleotides of the nucleic acid and are nonoverlapping, in order to avoid the formation of primer dimers. One of the primers will hybridize to one strand of the nucleic acid coding for the tumor-associated antigen, and the other primer will hybridize to the complementary strand in an arrangement which allows amplification of the nucleic acid coding for the tumor-associated antigen.

"Antisense" molecules or "antisense" nucleic acids may be used for regulating, in particular reducing, expression of a nucleic acid. The term "antisense molecule" or "antisense nucleic acid" refers according to the invention to an oligonucleotide which is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide or modified oligo-deoxyribonucleotide and which hybridizes under physiological conditions to DNA comprising a particular gene or to mRNA of said gene, thereby inhibiting transcription of said gene and/or translation of said mRNA. According to the invention, the "antisense molecule" also comprises a construct which contains a nucleic acid or a part thereof in reverse orientation with respect to its natural promoter. An antisense transcript of a nucleic acid or of a part thereof may form a duplex with the naturally occurring mRNA specifying the enzyme and thus prevent accumulation of or translation of the mRNA into the active enzyme. Another possibility is the use of ribozymes for inactivating a nucleic acid. Antisense oligonucleotides preferred according to the invention have a sequence of 6-50, in particular 10-30, 15-30 and 20-30, contiguous nucleotides of the target nucleic acid and preferably are fully complementary to the target nucleic acid or to a part thereof.

In preferred embodiments, the antisense oligonucleotide hybridizes with an N-terminal or 5' upstream site such as a translation initiation site, transcription initiation site or promoter site. In further embodiments, the antisense oligonucleotide hybridizes with a 3'untranslated region or mRNA splicing site.

In one embodiment, an oligonucleotide of the invention consists of ribonucleotides, deoxyribonucleotides or a combination thereof, with the 5' end of one nucleotide and the 3' end of another nucleotide being linked to

one another by a phosphodiester bond. These oligonucleotides may be synthesized in the conventional manner or produced recombinantly.

In preferred embodiments, an oligonucleotide of the invention is a "modified" oligonucleotide. Here, the oligonucleotide may be modified in very different ways, without impairing its ability to bind its target, in order to increase, for example, its stability or therapeutic efficacy. According to the invention, the term "modified oligonucleotide" means an oligonucleotide in which (i) at least two of its nucleotides are linked to one another by a synthetic internucleoside bond (i.e. an internucleoside bond which is not a phosphodiester bond) and/or (ii) a chemical group which is usually not found in nucleic acids is covalently linked to the oligonucleotide. Preferred synthetic internucleoside bonds are phosphorothioates, alkyl phosphonates, phosphorodithioates, phosphate esters, alkyl phosphonothioates, phosphoramidates, carbamates, carbonates, phosphate triesters, acetamidates, carboxymethyl esters and peptides.

The term "modified oligonucleotide" also comprises oligonucleotides having a covalently modified base and/or sugar. "Modified oligonucleotides" comprise, for example, oligonucleotides with sugar residues which are covalently bound to low molecular weight organic groups other than a hydroxyl group at the 3' position and a phosphate group at the 5' position. Modified oligonucleotides may comprise, for example, a 2'-O-alkylated ribose residue or another sugar instead of ribose, such as arabinose.

Preferably, the proteins and polypeptides described according to the invention have been isolated. The terms "isolated protein" or "isolated polypeptide" mean that the protein or polypeptide has been separated from

its natural environment. An isolated protein or polypeptide may be in an essentially purified state. The term "essentially purified" means that the protein or polypeptide is essentially free of other substances with which it is associated in nature or *in vivo*.

Such proteins and polypeptides may be used, for example, in producing antibodies and in an immunological or diagnostic assay or as therapeutics. Proteins and polypeptides described according to the invention may be isolated from biological samples such as tissue or cell homogenates and may also be expressed recombinantly in a multiplicity of pro- or eukaryotic expression systems.

For the purposes of the present invention, "derivatives" of a protein or polypeptide or of an amino acid sequence comprise amino acid insertion variants, amino acid deletion variants and/or amino acid substitution variants.

Amino acid insertion variants comprise amino- and/or carboxy-terminal fusions and also insertions of single or two or more amino acids in a particular amino acid sequence. In the case of amino acid sequence variants having an insertion, one or more amino acid residues are inserted into a particular site in an amino acid sequence, although random insertion with appropriate screening of the resulting product is also possible. Amino acid deletion variants are characterized by the removal of one or more amino acids from the sequence. Amino acid substitution variants are characterized by at least one residue in the sequence being removed and another residue being inserted in its place. Preference is given to the modifications being in positions in the amino acid sequence which are not conserved between homologous proteins or polypeptides. Preference is given to replacing amino acids with other ones having similar properties such as hydrophobicity,

hydrophilicity, electronegativity, volume of the side chain and the like (conservative substitution). Conservative substitutions, for example, relate to the exchange of one amino acid with another amino acid listed below in the same group as the amino acid to be substituted:

1. small aliphatic, nonpolar or slightly polar residues: Ala, Ser, Thr (Pro, Gly)
2. negatively charged residues and their amides: Asn, Asp, Glu, Gln
3. positively charged residues: His, Arg, Lys
4. large aliphatic, nonpolar residues: Met, Leu, Ile, Val (Cys)
5. large aromatic residues: Phe, Tyr, Trp.

Owing to their particular part in protein architecture, three residues are shown in brackets. Gly is the only residue without a side chain and thus imparts flexibility to the chain. Pro has an unusual geometry which greatly restricts the chain. Cys can form a disulfide bridge.

The amino acid variants described above may be readily prepared with the aid of known peptide synthesis techniques such as, for example, by solid phase synthesis (Merrifield, 1964) and similar methods or by recombinant DNA manipulation. Techniques for introducing substitution mutations at predetermined sites into DNA which has a known or partially known sequence are well known and comprise M13 mutagenesis, for example. The manipulation of DNA sequences for preparing proteins having substitutions, insertions or deletions, is described in detail in Sambrook et al. (1989), for example.

According to the invention, "derivatives" of proteins or polypeptides also comprise single or multiple substitutions, deletions and/or additions of any

molecules associated with the enzyme, such as carbohydrates, lipids and/or proteins or polypeptides. The term "derivative" also extends to all functional chemical equivalents of said proteins or polypeptides.

According to the invention, a part or fragment of a tumor-associated antigen has a functional property of the polypeptide from which it has been derived. Such functional properties comprise the interaction with antibodies, the interaction with other polypeptides or proteins, the selective binding of nucleic acids and an enzymatic activity. A particular property is the ability to form a complex with HLA and, where appropriate, generate an immune response. This immune response may be based on stimulating cytotoxic or T helper cells. A part or fragment of a tumor-associated antigen of the invention preferably comprises a sequence of at least 6, in particular at least 8, at least 10, at least 12, at least 15, at least 20, at least 30 or at least 50, consecutive amino acids of the tumor-associated antigen. A part or fragment of a tumor-associated antigen is preferably a part of the tumor-associated antigen which corresponds to the non-membrane portion, in particular the extracellular portion of the antigen or is comprised thereof.

A part or a fragment of a nucleic acid coding for a tumor-associated antigen relates according to the invention to the part of the nucleic acid, which codes at least for the tumor-associated antigen and/or for a part or a fragment of said tumor-associated antigen, as defined above. Preferably, a part or fragment of a nucleic acid coding for a tumor-associated antigen is that part which corresponds to the open reading frame, in particular as indicated in the sequence listing.

The isolation and identification of genes coding for tumor-associated antigens also make possible the

diagnosis of a disease characterized by expression of one or more tumor-associated antigens. These methods comprise determining one or more nucleic acids which code for a tumor-associated antigen and/or determining the encoded tumor-associated antigens and/or peptides derived therefrom. The nucleic acids may be determined in the conventional manner, including by polymerase chain reaction or hybridization with a labeled probe. Tumor-associated antigens or peptides derived therefrom may be determined by screening patient antisera with respect to recognizing the antigen and/or the peptides. They may also be determined by screening T cells of the patient for specificities for the corresponding tumor-associated antigen.

The present invention also enables proteins binding to tumor-associated antigens described herein to be isolated, including antibodies and cellular binding partners of said tumor-associated antigens.

According to the invention, particular embodiments ought to involve providing "dominant negative" polypeptides derived from tumor-associated antigens. A dominant negative polypeptide is an inactive protein variant which, by way of interacting with the cellular machinery, displaces an active protein from its interaction with the cellular machinery or which competes with the active protein, thereby reducing the effect of said active protein. For example, a dominant negative receptor which binds to a ligand but does not generate any signal as response to binding to the ligand can reduce the biological effect of said ligand. Similarly, a dominant negative catalytically inactive kinase which usually interacts with target proteins but does not phosphorylate said target proteins may reduce phosphorylation of said target proteins as response to a cellular signal. Similarly, a dominant negative transcription factor which binds to a promoter site in the control region of a gene but does not increase

transcription of said gene may reduce the effect of a normal transcription factor by occupying promoter binding sites, without increasing transcription.

The result of expression of a dominant negative polypeptide in a cell is a reduction in the function of active proteins. The skilled worker may prepare dominant negative variants of a protein, for example, by conventional mutagenesis methods and by evaluating the dominant negative effect of the variant polypeptide.

The invention also comprises substances such as polypeptides which bind to tumor-associated antigens. Such binding substances may be used, for example, in screening assays for detecting tumor-associated antigens and complexes of tumor-associated antigens with their binding partners and in a purification of said tumor-associated antigens and of complexes thereof with their binding partners. Such substances may also be used for inhibiting the activity of tumor-associated antigens, for example by binding to such antigens.

The invention therefore comprises binding substances such as, for example, antibodies or antibody fragments, which are capable of selectively binding to tumor-associated antigens. Antibodies comprise polyclonal and monoclonal antibodies which are produced in the conventional manner.

It is known that only a small part of an antibody molecule, the paratope, is involved in binding of the antibody to its epitope (cf. Clark, W.R. (1986), *The Experimental Foundations of Modern Immunology*, Wiley & Sons, Inc., New York; Roitt, I. (1991), *Essential Immunology*, 7th Edition, Blackwell Scientific Publications, Oxford). The pFc' and Fc regions are, for example, effectors of the complement cascade but are not involved in antigen binding. An antibody from which

the pFc′ region has been enzymatically removed or which has been produced without the pFc′ region, referred to as $F(ab')_2$ fragment, carries both antigen binding sites of a complete antibody. Similarly, an antibody from which the Fc region has been enzymatically removed or which has been produced without said Fc region, referred to Fab fragment, carries one antigen binding site of an intact antibody molecule. Furthermore, Fab fragments consist of a covalently bound light chain of an antibody and part of the heavy chain of said antibody, referred to as Fd. The Fd fragments are the main determinants of antibody specificity (a single Fd fragment can be associated with up to ten different light chains, without altering the specificity of the antibody) and Fd fragments, when isolated, retain the ability to bind to an epitope.

Located within the antigen-binding part of an antibody are complementary-determining regions (CDRs) which interact directly with the antigen epitope and framework regions (FRs) which maintain the tertiary structure of the paratope. Both the Fd fragment of the heavy chain and the light chain of IgG immunoglobulins contain four framework regions (FR1 to FR4) which are separated in each case by three complementary-determining regions (CDR1 to CDR3). The CDRs and, in particular, the CDR3 regions and, still more particularly, the CDR3 region of the heavy chain are responsible to a large extent for antibody specificity.

Non-CDR regions of a mammalian antibody are known to be able to be replaced by similar regions of antibodies with the same or a different specificity, with the specificity for the epitope of the original antibody being retained. This made possible the development of "humanized" antibodies in which nonhuman CDRs are covalently linked to human FR and/or Fc/pFc′ regions to produce a functional antibody.

WO 92/04381 for example, describes production and use of humanized murine RSV antibodies in which at least part of the murine FR regions have been replaced with FR regions of a human origin. Antibodies of this kind, including fragments of intact antibodies with antigen-binding capability, are often referred to as "chimeric" antibodies.

The invention also provides $F(ab')_2$, Fab, Fv, and Fd fragments of antibodies, chimeric antibodies, in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain-CDR3 regions have been replaced with homologous human or nonhuman sequences, chimeric $F(ab')_2$-fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain-CDR3 regions have been replaced with homologous human or nonhuman sequences, chimeric Fab-fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain-CDR3 regions have been replaced with homologous human or nonhuman sequences, and chimeric Fd-fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced with homologous human or nonhuman sequences. The invention also comprises "single-chain" antibodies.

Preferably, an antibody used according to the invention is directed against one of the sequences according to SEQ ID NOs: 2,6, 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, 66, 70, 74, 78, 82, 86, 90, 94, 98, 102, 106, 110, 114, 118, 122, 126, 130, 134, 138, 142, 146, 150, 154, 158, 162, 166, 170, 174, 176, 180, 184, 188, 192, 196, 200, 204, 208, 212, 216, 220, 224, 228, 232, 236, 240, 244, 248, 252, 256, 260, 264, 268, 270, 272, 274, 276, 278, 280 to 308, 310 of the sequence listing, a part or derivative thereof, in particular a sequence according to SEQ ID Nos: 281 to 308 of the sequence listing and/or may be obtained by immunization using these peptides.

The invention also comprises polypeptides which bind

specifically to tumor-associated antigens. Polypeptide binding substances of this kind may be provided, for example, by degenerate peptide libraries which may be prepared simply in solution in an immobilized form or as phage-display libraries. It is likewise possible to prepare combinatorial libraries of peptides with one or more amino acids. Libraries of peptoids and nonpeptidic synthetic residues may also be prepared.

Phage display may be particularly effective in identifying binding peptides of the invention. In this connection, for example, a phage library is prepared (using, for example, the M13, fd or lambda phages) which presents inserts of from 4 to about 80 amino acid residues in length. Phages are then selected which carry inserts which bind to the tumor-associated antigen. This process may be repeated via two or more cycles of a reselection of phages binding to the tumor-associated antigen. Repeated rounds result in a concentration of phages carrying particular sequences. An analysis of DNA sequences may be carried out in order to identify the sequences of the expressed polypeptides. The smallest linear portion of the sequence binding to the tumor-associated antigen may be determined. The "two-hybrid system" of yeast may also be used for identifying polypeptides which bind to a tumor-associated antigen. Tumor-associated antigens described according to the invention or fragments thereof may be used for screening peptide libraries, including phage-display libraries, in order to identify and select peptide binding partners of the tumor-associated antigens. Such molecules may be used, for example, for screening assays, purification protocols, for interference with the function of the tumor-associated antigen and for other purposes known to the skilled worker.

The antibodies described above and other binding molecules may be used, for example, for identifying

tissue which expresses a tumor-associated antigen. Antibodies may also be coupled to specific diagnostic substances for displaying cells and tissues expressing tumor-associated antigens. They may also be coupled to therapeutically useful substances. Diagnostic substances comprise, in a nonlimiting manner, barium sulfate, iocetamic acid, iopanoic acid, calcium ipodate, sodium diatrizoate, meglumine diatrizoate, metrizamide, sodium tyropanoate and radio diagnostic, including positron emitters such as fluorine-18 and carbon-11, gamma emitters such as iodine-123, technetium-99m, iodine-131 and indium-111, nuclides for nuclear magnetic resonance, such as fluorine and gadolinium. According to the invention, the term "therapeutically useful substance" means any therapeutic molecule which, as desired, is selectively guided to a cell which expresses one or more tumor-associated antigens, including anticancer agents, radioactive iodine-labeled compounds, toxins, cytostatic or cytolytic drugs, etc. Anticancer agents comprise, for example, aminoglutethimide, azathioprine, bleomycin sulfate, busulfan, carmustine, chlorambucil, cisplatin, cyclophosphamide, cyclosporine, cytarabidine, dacarbazine, dactinomycin, daunorubin, doxorubicin, taxol, etoposide, fluorouracil, interferon-α, lomustine, mercaptopurine, methotrexate, mitotane, procarbazine HCl, thioguanine, vinblastine sulfate and vincristine sulfate. Other anticancer agents are described, for example, in Goodman and Gilman, "The Pharmacological Basis of Therapeutics", 8th Edition, 1990, McGraw-Hill, Inc., in particular Chapter 52 (Antineoplastic Agents (Paul Calabresi and Bruce A. Chabner). Toxins may be proteins such as pokeweed antiviral protein, cholera toxin, pertussis toxin, ricin, gelonin, abrin, diphtheria exotoxin or *Pseudomonas* exotoxin. Toxin residues may also be high energy-emitting radionuclides such as cobalt-60.

The term "patient" means according to the invention a

human being, a nonhuman primate or another animal, in particular a mammal such as a cow, horse, pig, sheep, goat, dog, cat or a rodent such as a mouse and rat. In a particularly preferred embodiment, the patient is a human being.

According to the invention, the term "disease" refers to any pathological state in which tumor-associated antigens are expressed or abnormally expressed. "Abnormal expression" means according to the invention that expression is altered, preferably increased, compared to the state in a healthy individual. An increase in expression refers to an increase by at least 10%, in particular at least 20%, at least 50% or at least 100%. In one embodiment, the tumor-associated antigen is expressed only in tissue of a diseased individual, while expression in a healthy individual is repressed. One example of such a disease is cancer, in particular seminomas, melanomas, teratomas, gliomas, colon cancer, rectal cancer, kidney cancer, breast cancer, prostate cancer, cancer of the uterus, ovarian cancer, endometrial cancer, cancer of the esophagus, blood cancer, liver cancer, pancreatic cancer, skin cancer, brain cancer and lung cancer, lymphomas, and neuroblastomas. Examples for this are lung tumor, breast tumor, prostate tumor, colon tumor, renal cell carcinoma, cervical carcinoma, colon carcinoma and mamma carcinoma or metastases of the above cancer types or tumors.

According to the invention, a biological sample may be a tissue sample and/or a cellular sample and may be obtained in the conventional manner such as by tissue biopsy, including punch biopsy, and by taking blood, bronchial aspirate, urine, feces or other body fluids, for use in the various methods described herein.

According to the invention, the term "immunoreactive cell" means a cell which can mature into an immune cell

(such as B cell, T helper cell, or cytolytic T cell) with suitable stimulation. Immunoreactive cells comprise $CD34^+$ hematopoietic stem cells, immature and mature T cells and immature and mature B cells. If production of cytolytic or T helper cells recognizing a tumor-associated antigen is desired, the immunoreactive cell is contacted with a cell expressing a tumor-associated antigen under conditions which favor production, differentiation and/or selection of cytolytic T cells and of T helper cells. The differentiation of T cell precursors into a cytolytic T cell, when exposed to an antigen, is similar to clonal selection of the immune system.

Some therapeutic methods are based on a reaction of the immune system of a patient, which results in a lysis of antigen-presenting cells such as cancer cells which present one or more tumor-associated antigens. In this connection, for example autologous cytotoxic T lymphocytes specific for a complex of a tumor-associated antigen and an MHC molecule are administered to a patient having a cellular abnormality. The production of such cytotoxic T lymphocytes *in vitro* is known. An example of a method of differentiating T cells can be found in WO-A-96/33265. Generally, a sample containing cells such as blood cells is taken from the patient and the cells are contacted with a cell which presents the complex and which can cause propagation of cytotoxic T lymphocytes (e.g. dendritic cells). The target cell may be a transfected cell such as a COS cell. These transfected cells present the desired complex on their surface and, when contacted with cytotoxic T lymphocytes, stimulate propagation of the latter. The clonally expanded autologous cytotoxic T lymphocytes are then administered to the patient.

In another method of selecting antigen-specific cytotoxic T lymphocytes, fluorogenic tetramers of MHC class I molecule/peptide complexes are used for

detecting specific clones of cytotoxic T lymphocytes (Altman et al., *Science* 274:94-96, 1996; Dunbar et al., *Curr. Biol.* 8:413-416, 1998). Soluble MHC class I molecules are folded *in vitro* in the presence of β_2 microglobulin and a peptide antigen binding to said class I molecule. The MHC/peptide complexes are purified and then labeled with biotin. Tetramers are formed by mixing the biotinylated peptide-MHC complexes with labeled avidin (e.g. phycoerythrin) in a molar ratio of 4:1. Tetramers are then contacted with cytotoxic T lymphocytes such as peripheral blood or lymph nodes. The tetramers bind to cytotoxic T lymphocytes which recognize the peptide antigen/MHC class I complex. Cells which are bound to the tetramers may be sorted by fluorescence-controlled cell sorting to isolate reactive cytotoxic T lymphocytes. The isolated cytotoxic T lymphocytes may then be propagated *in vitro*.

In a therapeutic method referred to as adoptive transfer (Greenberg, *J. Immunol.* 136(5):1917, 1986; Riddel et al., *Science* 257:238, 1992; Lynch et al., *Eur. J. Immunol.* 21:1403-1410, 1991; Kast et al., *Cell* 59:603-614, 1989), cells presenting the desired complex (e.g. dendritic cells) are combined with cytotoxic T lymphocytes of the patient to be treated, resulting in a propagation of specific cytotoxic T lymphocytes. The propagated cytotoxic T lymphocytes are then administered to a patient having a cellular anomaly characterized by particular abnormal cells presenting the specific complex. The cytotoxic T lymphocytes then lyse the abnormal cells, thereby achieving a desired therapeutic effect.

Often, of the T cell repertoire of a patient, only T cells with low affinity for a specific complex of this kind can be propagated, since those with high affinity have been extinguished due to development of tolerance. An alternative here may be a transfer of the T cell

receptor itself. For this too, cells presenting the desired complex (e.g. dendritic cells) are combined with cytotoxic T lymphocytes of healthy individuals. This results in propagation of specific cytotoxic T lymphocytes with high affinity if the donor had no previous contact with the specific complex. The high affinity T cell receptor of these propagated specific T lymphocytes is cloned and can be transduced via gene transfer, for example using retroviral vectors, into T cells of other patients, as desired. Adoptive transfer is then carried out using these genetically altered T lymphocytes (Stanislawski et al., Nat Immunol. 2:962-70, 2001; Kessels et al., Nat Immunol. 2:957-61, 2001).

The therapeutic aspects above start out from the fact that at least some of the abnormal cells of the patient present a complex of a tumor-associated antigen and an HLA molecule. Such cells may be identified in a manner known per se. As soon as cells presenting the complex have been identified, they may be combined with a sample from the patient, which contains cytotoxic T lymphocytes. If the cytotoxic T lymphocytes lyse the cells presenting the complex, it can be assumed that a tumor-associated antigen is presented.

Adoptive transfer is not the only form of therapy which can be applied according to the invention. Cytotoxic T lymphocytes may also be generated *in vivo* in a manner known per se. One method uses nonproliferative cells expressing the complex. The cells used here will be those which usually express the complex, such as irradiated tumor cells or cells transfected with one or both genes necessary for presentation of the complex (i.e. the antigenic peptide and the presenting HLA molecule). Various cell types may be used. Furthermore, it is possible to use vectors which carry one or both of the genes of interest. Particular preference is given to viral or bacterial vectors. For example, nucleic acids coding for a tumor-associated antigen or

for a part thereof may be functionally linked to promoter and enhancer sequences which control expression of said tumor-associated antigen or a fragment thereof in particular tissues or cell types. The nucleic acid may be incorporated into an expression vector. Expression vectors may be nonmodified extrachromosomal nucleic acids, plasmids or viral genomes into which exogenous nucleic acids may be inserted. Nucleic acids coding for a tumor-associated antigen may also be inserted into a retroviral genome, thereby enabling the nucleic acid to be integrated into the genome of the target tissue or target cell. In these systems, a microorganism such as vaccinia virus, pox virus, Herpes simplex virus, retrovirus or adenovirus carries the gene of interest and de facto "infects" host cells. Another preferred form is the introduction of the tumor-associated antigen in the form of recombinant RNA which may be introduced into cells by liposomal transfer or by electroporation, for example. The resulting cells present the complex of interest and are recognized by autologous cytotoxic T lymphocytes which then propagate.

A similar effect can be achieved by combining the tumor-associated antigen or a fragment thereof with an adjuvant in order to make incorporation into antigen-presenting cells *in vivo* possible. The tumor-associated antigen or a fragment thereof may be represented as protein, as DNA (e.g. within a vector) or as RNA. The tumor-associated antigen is processed to produce a peptide partner for the HLA molecule, while a fragment thereof may be presented without the need for further processing. The latter is the case in particular, if these can bind to HLA molecules. Preference is given to administration forms in which the complete antigen is processed *in vivo* by a dendritic cell, since this may also produce T helper cell responses which are needed for an effective immune response (Ossendorp et al., *Immunol Lett.* 74:75-9, 2000; Ossendorp et al., *J. Exp.*

Med. 187:693-702, 1998). In general, it is possible to administer an effective amount of the tumor-associated antigen to a patient by intradermal injection, for example. However, injection may also be carried out intranodally into a lymph node (Maloy et al., *Proc Natl Acad Sci USA* 98:3299-303, 2001). It may also be carried out in combination with reagents which facilitate uptake into dendritic cells. Preferred tumor-associated antigens comprise those which react with allogenic cancer antisera or with T cells of many cancer patients. Of particular interest, however, are those against which no spontaneous immune responses pre-exist. Evidently, it is possible to induce against these immune responses which can lyse tumors (Keogh et al., *J. Immunol.* 167:787-96, 2001; Appella et al., *Biomed Pept Proteins Nucleic Acids* 1:177-84, 1995; Wentworth et al., *Mol Immunol.* 32:603-12, 1995).

The pharmaceutical compositions described according to the invention may also be used as vaccines for immunization. According to the invention, the terms "immunization" or "vaccination" mean an increase in or activation of an immune response to an antigen. It is possible to use animal models for testing an immunizing effect on cancer by using a tumor-associated antigen or a nucleic acid coding therefor. For example, human cancer cells may be introduced into a mouse to generate a tumor, and one or more nucleic acids coding for tumor-associated antigens may be administered. The effect on the cancer cells (for example reduction in tumor size) may be measured as a measure for the effectiveness of an immunization by the nucleic acid.

As part of the composition for an immunization, one or more tumor-associated antigens or stimulating fragments thereof are administered together with one or more adjuvants for inducing an immune response or for increasing an immune response. An adjuvant is a substance which is incorporated into the antigen or

administered together with the latter and which enhances the immune response. Adjuvants may enhance the immune response by providing an antigen reservoir (extracellularly or in macrophages), activating macrophages and stimulating particular lymphocytes. Adjuvants are known and comprise in a nonlimiting way monophosphoryl lipid A (MPL, SmithKline Beecham), saponin such as QS21 (SmithKline Beecham), DQS21 (SmithKline Beecham; WO 96/33739), QS7, QS17, QS18 and QS-L1 (So et al., Mol. Cells 7:178-186, 1997), incomplete Freund's adjuvant, complete Freund's adjuvant, vitamin E, montanide, alum, CpG oligonucleotides (cf. Krieg et al., Nature 374:546-9, 1995) and various water-in-oil emulsions prepared from biologically degradable oils such as squalene and/or tocopherol. Preferably, the peptides are administered in a mixture with DQS21/MPL. The ratio of DQS21 to MPL is typically about 1:10 to 10:1, preferably about 1:5 to 5:1 and in particular about 1:1. For administration to humans, a vaccine formulation typically contains DQS21 and MPL in a range from about 1 µg to about 100 µg.

Other substances which stimulate an immune response of the patient may also be administered. It is possible, for example, to use cytokines in a vaccination, owing to their regulatory properties on lymphocytes. Such cytokines comprise, for example, interleukin-12 (IL-12) which was shown to increase the protective actions of vaccines (cf. *Science* 268:1432-1434, 1995), GM-CSF and IL-18.

There are a number of compounds which enhance an immune response and which therefore may be used in a vaccination. Said compounds comprise costimulating molecules provided in the form of proteins or nucleic acids. Examples of such costimulating molecules are B7-1 and B7-2 (CD80 and CD86, respectively) which are expressed on dendritic cells (DC) and interact with the

CD28 molecule expressed on the T cells. This interaction provides a costimulation (signal 2) for an antigen/MHC/TCR-stimulated (signal 1) T cell, thereby enhancing propagation of said T cell and the effector function. B7 also interacts with CTLA4 (CD152) on T cells, and studies involving CTLA4 and B7 ligands demonstrate that B7-CTLA4 interaction can enhance antitumor immunity and CTL propagation (Zheng, P. et al., *Proc. Natl. Acad. Sci.* USA 95(11):6284-6289 (1998)).

B7 is typically not expressed on tumor cells so that these are no effective antigen-presenting cells (APCs) for T cells. Induction of B7 expression would enable tumor cells to stimulate more effectively propagation of cytotoxic T lymphocytes and an effector function. Costimulation by a combination of B7/IL-6/IL-12 revealed induction of IFN-gamma and Th1-cytokine profile in a T cell population, resulting in further enhanced T cell activity (Gajewski et al., *J. Immunol.* 154:5637-5648 (1995)).

A complete activation of cytotoxic T lymphocytes and a complete effector function require an involvement of T helper cells via interaction between the CD40 ligand on said T helper cells and the CD40 molecule expressed by dendritic cells (Ridge et al., *Nature* 393:474 (1998), Bennett et al., *Nature* 393:478 (1998), Schönberger et al., *Nature* 393:480 (1998)). The mechanism of this costimulating signal probably relates to the increase in B7 production and associated IL-6/IL-12 production by said dendritic cells (antigen-presenting cells). CD40-CD40L interaction thus complements the interaction of signal 1 (antigen/MHC-TCR) and signal 2 (B7-CD28).

The use of anti-CD40 antibodies for stimulating dendritic cells would be expected to directly enhance a response to tumor antigens which are usually outside

the range of an inflammatory response or which are presented by nonprofessional antigen-presenting cells (tumor cells). In these situations, T helper and B7-costimulating signals are not provided. This mechanism could be used in connection with therapies based on antigen-pulsed dendritic cells or in situations in which T helper epitopes have not been defined in known TRA precursors.

The invention also provides for administration of nucleic acids, polypeptides or peptides. Polypeptides and peptides may be administered in a manner known per se. In one embodiment, nucleic acids are administered by *ex vivo* methods, i.e. by removing cells from a patient, genetic modification of said cells in order to incorporate a tumor-associated antigen and reintroduction of the altered cells into the patient. This generally comprises introducing a functional copy of a gene into the cells of a patient *in vitro* and reintroducing the genetically altered cells into the patient. The functional copy of the gene is under the functional control of regulatory elements which allow the gene to be expressed in the genetically altered cells. Transfection and transduction methods are known to the skilled worker. The invention also provides for administering nucleic acids *in vivo* by using vectors such as viruses and target-controlled liposomes.

In a preferred embodiment, a viral vector for administering a nucleic acid coding for a tumor-associated antigen is selected from the group consisting of adenoviruses, adeno-associated viruses, pox viruses, including vaccinia virus and attenuated pox viruses, Semliki Forest virus, retroviruses, Sindbis virus and Ty virus-like particles. Particular preference is given to adenoviruses and retroviruses. The retroviruses are typically replication-deficient (i.e. they are incapable of generating infectious particles).

Various methods may be used in order to introduce according to the invention nucleic acids into cells *in vitro* or *in vivo*. Methods of this kind comprise transfection of nucleic acid $CaPO_4$ precipitates, transfection of nucleic acids associated with DEAE, transfection or infection with the above viruses carrying the nucleic acids of interest, liposome-mediated transfection, and the like. In particular embodiments, preference is given to directing the nucleic acid to particular cells. In such embodiments, a carrier used for administering a nucleic acid to a cell (e.g. a retrovirus or a liposome) may have a bound target control molecule. For example, a molecule such as an antibody specific for a surface membrane protein on the target cell or a ligand for a receptor on the target cell may be incorporated into or attached to the nucleic acid carrier. Preferred antibodies comprise antibodies which bind selectively a tumor-associated antigen. If administration of a nucleic acid via liposomes is desired, proteins binding to a surface membrane protein associated with endocytosis may be incorporated into the liposome formulation in order to make target control and/or uptake possible. Such proteins comprise capsid proteins or fragments thereof which are specific for a particular cell type, antibodies to proteins which are internalized, proteins addressing an intracellular site, and the like.

The therapeutic compositions of the invention may be administered in pharmaceutically compatible preparations. Such preparations may usually contain pharmaceutically compatible concentrations of salts, buffer substances, preservatives, carriers, supplementing immunity-enhancing substances such as adjuvants (e.g. CpG oligonucleotides) and cytokines and, where appropriate, other therapeutically active compounds.

The therapeutically active compounds of the invention may be administered via any conventional route, including by injection or infusion. The administration may be carried out, for example, orally, intravenously, intraperitonealy, intramuscularly, subcutaneously or transdermally. Preferably, antibodies are therapeutically administered by way of a lung aerosol. Antisense nucleic acids are preferably administered by slow intravenous administration.

The compositions of the invention are administered in effective amounts. An "effective amount" refers to the amount which achieves a desired reaction or a desired effect alone or together with further doses. In the case of treatment of a particular disease or of a particular condition characterized by expression of one or more tumor-associated antigens, the desired reaction relates to inhibition of the course of the disease. This comprises slowing down the progress of the disease and, in particular, interrupting the progress of the disease. The desired reaction in a treatment of a disease or of a condition may also be delay of the onset or a prevention of the onset of said disease or said condition.

An effective amount of a composition of the invention will depend on the condition to be treated, the severeness of the disease, the individual parameters of the patient, including age, physiological condition, size and weight, the duration of treatment, the type of an accompanying therapy (if present), the specific route of administration and similar factors.

The pharmaceutical compositions of the invention are preferably sterile and contain an effective amount of the therapeutically active substance to generate the desired reaction or the desired effect.

The doses administered of the compositions of the

invention may depend on various parameters such as the type of administration, the condition of the patient, the desired period of administration, etc. In the case that a reaction in a patient is insufficient with an initial dose, higher doses (or effectively higher doses achieved by a different, more localized route of administration) may be used.

Generally, doses of the tumor-associated antigen of from 1 ng to 1 mg, preferably from 10 ng to 100 µg, are formulated and administered for a treatment or for generating or increasing an immune response. If the administration of nucleic acids (DNA and RNA) coding for tumor-associated antigens is desired, doses of from 1 ng to 0.1 mg are formulated and administered.

The pharmaceutical compositions of the invention are generally administered in pharmaceutically compatible amounts and in pharmaceutically compatible compositions. The term "pharmaceutically compatible" refers to a nontoxic material which does not interact with the action of the active component of the pharmaceutical composition. Preparations of this kind may usually contain salts, buffer substances, preservatives, carriers and, where appropriate, other therapeutically active compounds. When used in medicine, the salts should be pharmaceutically compatible. However, salts which are not pharmaceutically compatible may used for preparing pharmaceutically compatible salts and are included in the invention. Pharmacologically and pharmaceutically compatible salts of this kind comprise in a nonlimiting way those prepared from the following acids: hydrochloric, hydrobromic, sulfuric, nitric, phosphoric, maleic, acetic, salicylic, citric, formic, malonic, succinic acids, and the like. Pharmaceutically compatible salts may also be prepared as alkali metal salts or alkaline earth metal salts, such as sodium salts, potassium salts or calcium salts.

A pharmaceutical composition of the invention may comprise a pharmaceutically compatible carrier. According to the invention, the term "pharmaceutically compatible carrier" refers to one or more compatible solid or liquid fillers, diluents or encapsulating substances, which are suitable for administration to humans. The term "carrier" refers to an organic or inorganic component, of a natural or synthetic nature, in which the active component is combined in order to facilitate application. The components of the pharmaceutical composition of the invention are usually such that no interaction occurs which substantially impairs the desired pharmaceutical efficacy.

The pharmaceutical compositions of the invention may contain suitable buffer substances such as acetic acid in a salt, citric acid in a salt, boric acid in a salt and phosphoric acid in a salt.

The pharmaceutical compositions may, where appropriate, also contain suitable preservatives such as benzalkonium chloride, chlorobutanol, paraben and thimerosal.

The pharmaceutical compositions are usually provided in a uniform dosage form and may be prepared in a manner known per se. Pharmaceutical compositions of the invention may be in the form of capsules, tablets, lozenges, solutions, suspensions, syrups, elixir or in the form of an emulsion, for example.

Compositions suitable for parenteral administration usually comprise a sterile aqueous or nonaqueous preparation of the active compound, which is preferably isotonic to the blood of the recipient. Examples of compatible carriers and solvents are Ringer solution and isotonic sodium chloride solution. In addition,

usually sterile, fixed oils are used as solution or suspension medium.

The present invention is described in detail by the figures and examples below, which are used only for illustration purposes and are not meant to be limiting. Owing to the description and the examples, further embodiments which are likewise included in the invention are accessible to the skilled worker.

Figures:

Fig. 1: qPCR analysis of SEQ ID NO: 1 in melanomas
Quantitative expression analysis of SEQ ID NO: 1 in healthy skin tissue, in testis and in melanomas. Logarithmic representation of relative expression (-fold activation).

Fig. 2: Conventional RT-PCR analysis of SEQ ID NO: 1 in melanomas
RT-PCR expression analysis of SEQ ID NO: 1 in melanomas (n = 14) and melanoma cell lines (n = 4) in comparison with healthy skin (n = 4) and with testis (n = 3).

Fig. 3: qPCR analysis of SEQ ID NO: 5 in healthy tissue and in tumor samples
Quantitative expression analysis of SEQ ID NO: 5 in normal tissues (left-hand side) and in various tumors (pools consisting of in each case 3-5 individual samples, right-hand side). **A** Logarithmic representation of relative expression (-fold activation). **B** Image after gel-electrophoretic fractionation of the amplified fragments.

Fig. 4: Detailed analysis of SEQ ID NO: 5-specific expression
A Quantitative expression analysis of SEQ ID NO: 5 in various ENT, renal and uterine tumors in comparison with expression in the corresponding normal tissues.

Logarithmic representation. **B** Image after gel-electrophoretic fractionation of the amplified fragments.

Fig. 5: Northern blot analysis with a SEQ ID NO: 5-specific sequence
Hybridization of a DIG-labeled DNA probe, prepared by PCR amplification using the primers according to SEQ ID NO: 7 and 8, with testis-specific RNA. Lane 1: 2 µg of testis-specific RNA; lane 2: 1 µg of testis-specific RNA.

Fig. 6: qPCR analysis of LOC203413
Quantitative expression analysis of LOC203413 in healthy tissues (left) and in tumor samples (pools consisting of in each case 3-5 individual samples, right). **A** Logarithmic representation of expression (-fold activation). **B** Result after gel-electrophoretic fractionation.

Fig. 7: Detailed analysis of LOC203413-specific expression in gastric carcinomas
Quantitative expression analysis of LOC203413 in various gastric tumor samples (n = 10) in comparison with expression in healthy stomach (n = 6). **A** Linear representation of relative expression. **B** Image after gel-electrophoretic fractionation of the amplicons.

Fig. 8: qPCR analysis of LOC90625-specific expression
Quantitative expression analysis of LOC90625 in normal tissues (left) and tumor tissues (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation).

Fig. 9: Detailed analysis of LOC90652-specific expression in various types of tumors
Quantitative expression analysis of LOC90625 in samples of carcinomas of the esophagus (n = 8), pancreas

(n = 5) and prostate (n = 10) in comparison with the respective healthy tissue (n = 3/4); logarithmic representation of relative expression (-fold activation).

Fig. 10: qRT-PCR analysis of FAM26A in various types of tumors

Quantitative RT-PCR expression analysis of FAM26A in carcinomas of the **A** ovary, **B** stomach, esophagus, pancreas and liver, in comparison with the respective healthy tissue. Linear representation of relative expression (-fold activation).

Fig. 11: Characterization of FAM26A-specific antibodies

Western blot analysis of the antisera generated by immunization with a peptide of SEQ ID NO: 291 **(A)** and SEQ ID NO: 292 **(B)**. Extracts of CHO cells were analyzed after transfection with in each case epitope-specific (A 1, 3; B 2, 4) or in each case epitope-unspecific (A 2, 4; B 1, 3) plasmids. The arrow indicates the specific fragments.

Fig. 12: Analysis of the FAM26A protein in tumors

Detection of FAM26A in cervical, ovarian and pancreatic tumors by means of FAM26A-specific antibodies (SEQ ID NO: 292).

Fig. 13: Analysis of the FAM26A protein in cell lines

Analysis of the FAM26A protein in cell lines with the aid of SEQ ID NO: 291-specific antibodies. **A** Western blot analysis with preimmune serum as specificity control (lanes 1-5) and FAM26A-specific antibodies. **B** Immunofluorescence analysis of SW480 cells.

Fig. 14: Immunohistochemical detection of FAM26A in testis

Immunohistochemical analysis of the FAM26A protein in healthy testis with the aid of SEQ ID NO: 292-specific antiserum in different dilutions **(A-C)**.

Fig. 15: Immunohistochemical analysis of FAM26A in tumors
Immunohistochemical analysis of the FAM26A protein in carcinoma samples (40-fold magnification, 1:300 dilution) with the aid of the SEQ ID NO: 292-specific antiserum. **A** Ovarian papillary cystadenocarcinoma. **B** Plate epithelial carcinoma of the cervix.

Fig. 16: qRT-PCR analysis of SEMA5B-specific expression
Quantitative expression analysis of SEMA5B in normal tissues (left) and tumor samples (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation).

Fig. 17: Detailed analysis of SEMA5B-specific expression in renal cell carcinoma samples
Quantitative expression analysis of SEMA5B in **A** renal cell carcinoma samples (n = 12) in comparison with healthy renal tissue (N = 3) and in **B** mammary carcinomas (N = 12) in comparison with healthy breast tissue (N = 3); logarithmic representation of relative expression (-fold activation).

Fig. 18: qRT-PCR analysis of GJB5-specific expression
Quantitative expression analysis of GJB5 in healthy tissue samples (left) and carcinomas (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation).

Fig. 19: Detailed analysis of GJB5-specific expression in various types of tumors
Quantitative expression analysis of GJB5 in **A** colon carcinoma samples (n = 12), **B** esophageal tumors (n = 8), **C** gastric carcinomas (n = 10) and **D** pancreatic tumors (n = 5) in comparison with in each case healthy tissue samples; logarithmic (A, C) or linear (B, D) representation of relative expression (-fold

activation).

Fig. 20: qRT-PCR analysis of KLK5-specific expression
Quantitative expression analysis of KLK5 in healthy tissue samples (left) and tumors (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation).

Fig. 21: Detailed analysis of KLK5-specific expression in various types of tumors
Quantitative expression analysis of KLK5 in esophageal tumors (n = 8), in ENT carcinomas (n = 5) and in cervical tumors (n = 4) in comparison with the respective healthy tissue samples; logarithmic representation of relative expression (-fold activation).

Fig. 22: qRT-PCR analysis of LOC352765-specific expression
Quantitative expression analysis of LOC352765 in healthy tissue samples (left) and tumors (pools consisting of in each case 3-5 individual samples; right). Logarithmic representation of relative expression (-fold activation).

Fig. 23: Detailed analysis of LOC352765-specific expression in various types of tumors
Quantitative expression analysis of LOC352765 in colon carcinomas (n = 8), in mammary carcinomas (n = 5) and in ENT tumors (n = 4) in comparison with respective healthy tissue samples; logarithmic representation of relative expression (-fold activation).

Fig. 24: qRT-PCR analysis of SVCT1-specific expression
Quantitative expression analysis of SVCT1 in healthy tissue samples (left) and tumors (pools consisting of in each case 3-5 individual samples; right). Logarithmic representation of relative expression

(-fold activation).

Fig. 25: Detailed analysis of SVCT1-specific expression in various types of tumors

Quantitative expression analysis of SVCT1 in **A** kidney carcinomas (n = 8), **B** esophageal tumors (n = 5) and ENT tumors (n = 4) in comparison with in each case healthy tissue samples; logarithmic representation of relative expression (-fold activation).

Fig. 26: qRT-PCR analysis of LOC199953-specific expression in renal cell carcinomas and in ENT tumors

Quantitative expression analysis of LOC199953 in renal cell carcinomas (n = 12) and ENT tumors (n = 5) in comparison with healthy kidney- and skin-specific tissue samples; linear representation of relative expression (-fold activation).

Fig. 27: qRT-PCR analysis of TMEM31-specific expression

Quantitative expression analysis of TMEM31 in healthy tissue samples (left) and tumors (pools consisting of in each case 3-5 individual samples; right). Logarithmic representation of relative expression (-fold activation).

Fig. 28: Detailed analysis of TMEM31-specific expression in various types of tumors

Quantitative expression analysis of TMEM31 in **A** gastric carcinomas (n = 10) and **B** mammary carcinomas (n = 12) in comparison with in each case healthy tissue samples; logarithmic representation of relative expression (-fold activation).

Fig. 29: qRT-PCR analysis of FLJ25132-specific expression in ovarian tumors and in prostate carcinomas

Quantitative expression analysis of FLJ25132 in ovarian tumors (n = 8) and in prostate carcinomas (n = 10) in comparison with in each case healthy tissue samples; linear representation of relative expression (-fold

activation).

Fig. 30: qRT-PCR analysis of SEQ ID NO: 57-specific expression

Quantitative expression analysis of SEQ ID NO: 57 in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation).

Fig. 31: Detailed analysis of SEQ ID NO: 57-specific expression in various types of tumors

Quantitative expression analysis of SEQ ID NO: 57 in **A** esophageal tumors (n = 8), **B** liver carcinomas (n = 8), **C** kidney carcinomas and **D** cervical and ENT tumors in comparison with in each case healthy tissue samples; linear (A, C, D) or logarithmic (B) representation of relative expression (-fold activation).

Fig. 32: qRT-PCR analysis of LOC119395-specific expression

Quantitative expression analysis of LOC119395 in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation).

Fig. 33: Detailed analysis of LOC119395-specific expression in various types of tumors

Quantitative expression analysis of LOC119395 in **A** breast tumors (n = 12), **B** esophageal carcinomas (n = 8) and **C** colon and gastric carcinomas, in comparison with in each case healthy tissue samples; logarithmic representation of relative expression (-fold activation).

Fig. 34: qRT-PCR analysis of LOC121838-specific expression

A Quantitative analysis of LOC121838-specific expres-

sion in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation). **B** Detailed analysis of LOC121838-specific RNA in ovarian tissues, logarithmic representation.

Fig. 35: qRT-PCR analysis of LOC221103-specific expression

Quantitative expression analysis of LOC221103-RNA in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation).

Fig. 36: Detailed qRT-PCR analysis of LOC221103-specific expression in liver samples

Quantitative expression analysis of LOC221103-RNA in liver tumors (n = 8) and in a healthy liver sample. Linear representation of relative expression (-fold activation).

Fig. 37: qRT-PCR analysis of LOC338579-specific expression

Quantitative expression analysis of LOC338579-specific RNA in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Logarithmic representation of relative expression (-fold activation).

Fig. 38: qRT-PCR analysis of LOC90342-specific expression

Quantitative expression analysis of LOC90342-specific RNA in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Logarithmic representation of relative expression (-fold activation).

Fig. 39: qRT-PCR analysis of LRFN1-specific expression
Quantitative expression analysis of LRFN1-specific RNA in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Logarithmic representation of relative expression (-fold activation).

Fig. 40: qRT-PCR analysis of LOC285916-specific expression
A Quantitative analysis of LOC285916-specific expression in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation). **B** Detailed analysis of LOC285916-specific RNA in kidney tissues and in ENT tumors, logarithmic representation.

Fig. 41: qRT-PCR analysis of MGC71744-specific expression
A Quantitative analysis of MGC71744-specific expression in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation). **B** Detailed analysis of MGC71744-specific RNA in various kidney tissues, logarithmic representation.

Fig. 42: qRT-PCR analysis of LOC342982-specific expression
Quantitative expression analysis of LOC342982-specific RNA in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Logarithmic representation of relative expression (-fold activation).

Fig. 43: qRT-PCR analysis of LOC343169-specific expression
A Quantitative analysis of LOC343169-specific expression in healthy tissue samples (left) and in tumors

(pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation). **B** Detailed analysis of LOC343169-specific RNA in various ovarian tissues, logarithmic representation.

Fig. 44: qRT-PCR analysis of LOC340204-specific expression

A Quantitative analysis of LOC340204-specific expression in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation). **B** Gel image of selected tissue samples after gel-electrophoretic fractionation.

Fig. 45: qRT-PCR analysis of LOC340067-specific expression

Quantitative expression analysis of LOC340067-specific RNA in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Logarithmic representation of relative expression (-fold activation).

Fig. 46: qRT-PCR analysis of LOC342780-specific expression

Quantitative expression analysis of LOC342780-specific RNA in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Logarithmic representation of relative expression (-fold activation).

Fig. 47: qRT-PCR analysis of LOC339511-specific expression

A Quantitative analysis of LOC339511-specific expression in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation). **B** Detailed analysis of LOC339511-specific RNA in various liver-specific

tissues; linear representation.

Fig. 48: qRT-PCR analysis of C14orf37-specific expression
Quantitative expression analysis of C14orf37 in healthy tissue samples (left) and in tumors (pools consisting of in each case 3-5 individual samples; right). Linear representation of relative expression (-fold activation).

Fig. 49: qRT-PCR analysis of ATP1A4-specific expression
A Quantitative expression analysis of ATP1A4 in healthy tissue samples and in tumors (pools consisting of in each case 3-5 individual samples). Logarithmic representation of relative expression (-fold activation). **B** Detailed analysis of ATP1A4-specific RNA in various breast-specific tissues; logarithmic representation.

Examples:

Materials and methods

The terms "*in silico*" and "electronic" refer solely to the utilization of methods based on databases, which may also be used to simulate laboratory experimental processes.

Unless expressly defined otherwise, all other terms and expressions are used so as to be understood by the skilled worker. The techniques and methods mentioned are carried out in a manner known per se and are described, for example, in Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd edition (1989), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. All methods including the use of kits and reagents are carried out according to the manufacturers' information.

Example 1:
Data mining-based strategy for identifying tumor-associated antigens

According to the invention, public human protein and nucleic acid databases were screened with regard to cancer-specific antigens accessible on the cell surface. The definition of the screening criteria required therefor, together with high throughput methods for analyzing, if possible, all proteins, formed the central component of this strategy.

The starting point consisted of the potential genes, predicted mainly by the human genome project, which have been deposited as solely exemplary protein (XP) or mRNA (XM) entries in the RefSeq database (Pruitt et al., *Trends Genet.* Jan; 16(1):44-47, 2000) of the National Center for Biotechnology Information (NCBI). In another approach, the validated protein entries (NP) and, respectively, the corresponding mRNAs (NM) of the

same database were also analyzed in the same manner. Following the fundamental principle of (hypothetical) gene → mRNA → protein, the proteins were first studied for the presence of transmembrane domains by combining a plurality of prediction programs for protein analysis. A total of 19 544 entries of the human XP fraction of the RefSeq database were analyzed, with 2025 hypothetical proteins satisfying said screening criteria. The human NP fraction provided a total of 19 110 entries with a proportion of 4634 filtered proteins.

The corresponding mRNA of each of these 2025 and 4634 proteins, respectively, was then subjected to a homology search in the EST database (Boguski et al., *Nat. Genet.* 4(4):332-333, 1993) of the NCBI with the aid of the BLAST algorithm (Altschul et al., *Nucleic Acids Res.* 25:3389-3402, 1997). The screening criteria in this search were set to stringent. A total of 1270 hypothetic mRNAs scored at least one hit in the EST database, with the number of hits exceeding 1000 in some cases.

Subsequently, the tissue-specific origin of the underlying cDNA library as well as the name of the library were determined for each of these valid hits. The tissues resulting therefrom were divided into 4 different groups ranging from dispensable organs (group 3) to absolutely essential organs (group 0). Another group, group 4, consisted of any samples obtained from cancer tissue. The distribution of hits to the five groups was recorded in a table which was sorted according to the best ratio of the sum of groups 3 and 4 to the sum of groups 0-2. Those mRNAs whose EST hits originated exclusively from cancer tissue reached a top position, followed by those which can additionally be found also in tissues of dispensable organs of group 3.

Since the transcripts determined in the first approach and the corresponding proteins are firstly hypothetic constructs, further screening criteria were used with the intention to prove the real existence of the mRNAs and consequently also of the proteins. For this purpose, each mRNA was compared to the predicted gene locus. Only those transcripts which have at least one splicing process, i.e. which spread over at least 2 exons, were used for more detailed analyses.

Sequential application of all the filters mentioned led to the tumor-associated antigens of the invention which can be considered extracellularly accessible, owing to a predicted transmembrane domain and the topology related thereto. The expression profile derived from the EST data indicates, in all cases, cancer-specific expression which may at most extend only to dispensable organs.

Example 2:

Strategy of validating the tumor-associated antigens identified by *in silico* analysis

In order to utilize the targets for immunotherapeutic purposes (antibody therapy by means of monoclonal antibodies, vaccination, T-cell receptor-mediated therapeutic approaches; cf. EP-B-0 879 282) or other targeted approaches (small compounds, siRNA etc.) in cancer therapy as well as for diagnostic problems, the validation of the targets identified according to the invention is of central importance. In this connection, validation is carried out by expression analysis at both RNA and protein levels.

1. Examination of RNA expression

The identified tumor antigens are first validated with the aid of RNA which is obtained from various tissues or from tissue-specific cell lines. Since the differential expression pattern of healthy tissue in

comparison with tumor tissue is of decisive importance for the subsequent therapeutic application, the target genes are preferably characterized with the aid of these tissue samples.

Total RNA is isolated from native tissue samples or from tumor cell lines by standard methods of molecular biology. Said isolation may be carried out, for example, with the aid of the RNeasy Maxi kit (Qiagen, Cat. No. 75162) according to the manufacturer's instructions. This isolation method is based on the use of chaotropic reagent guanidinium isothiocyanate. Alternatively, acidic phenol can be used for isolation (Chomczynski & Sacchi, Anal. Biochem. 162: 156-159, 1987). After the tissue has been worked up by means of guanidinium isothiocyanate, RNA is extracted with acidic phenol, subsequently precipitated with isopropanol and taken up in DEPC-treated water.

2-4 µg of the RNA isolated in this way are subsequently transcribed into cDNA, for example by means of Superscript II (Invitrogen) according to the manufacturer's protocol. cDNA synthesis is primed with the aid of random hexamers (e.g. Roche Diagnostics) according to standard protocols of the relevant manufacturer. For quality control, the cDNAs are amplified over 30 cycles, using primers specific for the p53 gene which is expressed only lowly. Only p53-positive cDNA samples will be used for the subsequent reaction steps.

The targets are analyzed in detail by carrying out an expression analysis by means of PCR or quantitative PCR (qPCR) on the basis of a cDNA archive which has been isolated from various normal and tumor tissues and from tumor cell lines. For this purpose, 0.5 µl of cDNA of the above reaction mixture is amplified by a DNA polymerase (e.g. 1 U of HotStarTaq DNA polymerase, Qiagen) according to the protocols of the particular

manufacturer (total volume of the reaction mixture: 25-50 µl). Aside from said polymerase, the amplification mixture comprises 0.3 mM dNTPs, reaction buffer (final concentration 1 ×, depending on the manufacturer of the DNA polymerase) and in each case 0.3 mM gene-specific forward and reverse primers.

The specific primers of the target gene are, as far as possible, selected in such a way that they are located in two different exons so that genomic contaminations do not lead to false-positive results. In a non-quantitative end point PCR, the cDNA is typically incubated at 95°C for 15 minutes in order to denature the DNA and to activate the Hot-Start enzyme. Subsequently the DNA is amplified over 35 cycles (1 min at 95°C, 1 min at the primer-specific hybridization temperature (approx. 55-65°C), 1 min at 72°C to elongate the amplicons). Subsequently, 10 µl of the PCR mixture are applied to agarose gels and fractionated in the electric field. The DNA is made visible in the gels by staining with ethidium bromide and the PCR result is documented by way of a photograph.

As an alternative to conventional PCR, expression of a target gene may also be analyzed by quantitative real time PCR. Meanwhile various analytical systems are available for this analysis, of which the best known ones are the ABI PRISM sequence detection system (TaqMan, Applied Biosystems), the iCycler (Biorad) and the Light cycler (Roche Diagnostics). As described above, a specific PCR mixture is subjected to a run in the real time instruments. By adding a DNA-intercalating dye (e.g. ethidium bromide, CybrGreen), the newly synthesized DNA is made visible by specific light excitation (according to the dye manufacturers' information). A multiplicity of points measured during amplification enables the entire process to be monitored and the nucleic acid concentration of the target gene to be determined quantitatively. The PCR

mixture is normalized by measuring a housekeeping gene (e.g. 18S RNA, β-actin). Alternative strategies via fluorescently labeled DNA probes likewise allow quantitative determination of the target gene of a specific tissue sample (see TaqMan applications from Applied Biosystems).

2. Cloning

The complete target gene which is required for further characterization of the tumor antigen is cloned according to common molecular-biological methods (e.g. in "Current Protocols in Molecular Biology", John Wiley & Sons Ltd., Wiley InterScience). In order to clone the target gene or to analyze its sequence, said gene is first amplified by a DNA polymerase having a proof reading function (e.g. pfu, Roche Diagnostics). The amplicon is then ligated by standard methods into a cloning vector. Positive clones are identified by sequence analysis and subsequently characterized with the aid of prediction programs and known algorithms.

3. Prediction of the protein

Many of the genes found according to the invention (in particular those from the RefSeq XM domain) are newly discovered genes which require cloning of the full-length gene, determination of the open reading frame and deduction and analysis of the protein sequence.
In order to clone the full-length sequence, we used common protocols for the rapid amplification of cDNA ends and the screening of cDNA expression libraries with gene-specific probes (*Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd edition (1989), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.*).
After assembling the fragments found in this way, potential open reading frames (ORF) were predicted using common prediction programs. Since the position of

the PolyA tail and of polyadenylation motifs predetermines the orientation of the potential gene product, only the 3 reading frames of that particular orientation remain out of a possible 6 reading frames. The former often yield only one sufficiently large open reading frame which may code for a protein, while the other reading frames have too many stop codons and would not code for any realistic protein. In the case of alternative open reading frames, identification of the authentic ORF is assisted by taking into account the Kozak criteria for optimal transcription initiation and by analyzing the deduced protein sequences which may arise. Said ORF is further verified by generating immune sera against proteins deduced from the potential ORFs and analyzing said immune sera for recognition of a real protein in tissues and cell lines.

4. Production of antibodies

The tumor-associated antigens identified according to the invention are characterized, for example, by using antibodies. The invention further comprises the diagnostic or therapeutic use of antibodies. Antibodies may recognize proteins in the native and/or denatured state (Anderson et al., *J. Immunol.* 143: 1899-1904, 1989; Gardsvoll, *J. Immunol. Methods* 234: 107-116, 2000; Kayyem et al., *Eur. J. Biochem.* 208: 1-8, 1992; Spiller et al., *J. Immunol. Methods* 224: 51-60, 1999).

Antisera comprising specific antibodies which specifically bind to the target protein may be prepared by various standard methods; cf., for example, "Monoclonal Antibodies: A Practical Approach" by Phillip Shepherd, Christopher Dean ISBN 0-19-963722-9, "Antibodies: A Laboratory Manual" by Ed Harlow, David Lane ISBN: 0879693142 and "Using Antibodies: A Laboratory Manual: Portable Protocol NO" by Edward Harlow, David Lane, Ed Harlow ISBN: 0879695447. It is also possible here to generate affine and

specific antibodies which recognize complex membrane proteins in their native form (Azorsa et al., *J. Immunol. Methods* 229: 35-48, 1999; Anderson et al., *J. Immunol.* 143: 1899-1904, 1989; Gardsvoll, *J. Immunol. Methods.* 234: 107-116, 2000). This is especially important in the preparation of antibodies which are intended to be used therapeutically but also for many diagnostic applications. For this purpose, both the complete protein and extracellular partial sequences may be used for immunization.

Immunization and production of polyclonal antibodies

A species (e.g. rabbits, mice) is immunized by a first injection of the desired target protein. The immune response of the animal to the immunogen can be enhanced by a second or third immunization within a defined period of time (approx. 2-4 weeks after the previous immunization). Blood is taken from said animals and immune sera obtained, again after various defined time intervals (1st bleeding after 4 weeks, then every 2-3 weeks, up to 5 takings). The immune sera taken in this way comprise polyclonal antibodies which may be used to detect and characterize the target protein in Western blotting, by flow cytometry, immunofluorescence or immunohistochemistry.

The animals are usually immunized by any of four well-established methods, with other methods also in existence. The immunization may be carried out using peptides specific for the target protein, using the complete protein, using extracellular partial sequences of a protein which can be identified experimentally or via prediction programs. Since the prediction programs do not always work perfectly, it is also possible to employ two domains separated from one another by a transmembrane domain. In this case, one of the two domains has to be extracellular, which may then be proved experimentally (see below).

(1) In the first case, peptides (length: 8-12 amino acids) are synthesized by *in vitro* methods (possibly carried out by a commercial service), and said peptides are used for immunization. Normally 3 immunizations are carried out (e.g. with a concentration of 5-100 µg/immunization). The immunization may also be carried out by commercial service providers.

(2) Alternatively, immunization may be carried out using recombinant proteins. For this purpose, the cloned DNA of the target gene is cloned into an expression vector and the target protein is synthesized, for example, cell-free *in vitro*, in bacteria (e.g. *E. coli*), in yeast (e.g. *S. pombe*), in insect cells or in mammalian cells, according to the conditions of the particular manufacturer (e.g. Roche Diagnostics, Invitrogen, Clontech, Qiagen). It is also possible to synthesize the target protein with the aid of viral expression systems (e.g. baculovirus, vacciniavirus, adenovirus). After it has been synthesized in one of said systems, the target protein is purified, normally by employing chromatographic methods. In this context, it is also possible to use for immunization proteins which have a molecular anchor as an aid for purification (e.g. His tag, Qiagen; FLAG tag, Roche Diagnostics; GST fusion proteins). A multiplicity of protocols can be found, for example, in "Current Protocols in Molecular Biology", John Wiley & Sons Ltd., Wiley InterScience. After the target protein has been purified, an immunization is carried out as described above.

(3) If a cell line is available which synthesizes the desired protein endogenously, it is also possible to use this cell line directly for preparing the specific antiserum. In this case, immunization is carried out by 1-3 injections with in each case approx. 1-5 × 10^7 cells.

(4) The immunization may also be carried out by injecting DNA (DNA immunization). For this purpose, the target gene is first cloned into an expression vector so that the target sequence is under the control of a strong eukaryotic promoter (e.g. CMV promoter). Subsequently, DNA (e.g. 1-10 µg per injection) is transferred as immunogen using a gene gun into capillary regions with a strong blood flow in an organism (e.g. mouse, rabbit). The transferred DNA is taken up by the animal's cells, the target gene is expressed, and the animal finally develops an immune response to the target protein (Jung et al., *Mol. Cells* 12: 41-49, 2001; Kasinrerk et al., *Hybrid Hybridomics* 21: 287-293, 2002).

Production of monoclonal antibodies

Monoclonal antibodies are traditionally produced with the aid of the hybridoma technology (technical details: see "Monoclonal Antibodies: A Practical Approach" by Philip Shepherd, Christopher Dean ISBN 0-19-963722-9; "Antibodies: A Laboratory Manual" by Ed Harlow, David Lane ISBN: 0879693142, "Using Antibodies: A Laboratory Manual: Portable Protocol NO" by Edward Harlow, David Lane, Ed Harlow ISBN: 0879695447). A new method which is also used is the "SLAM" technology. Here, B cells are isolated from whole blood and the cells are made monoclonal. Subsequently the supernatant of the isolated B cell is analyzed for its antibody specificity. In contrast to the hybridoma

technology, the variable region of the antibody gene is then amplified by single-cell PCR and cloned into a suitable vector. In this manner production of monoclonal antibodies is accelerated (de Wildt et al., J. Immunol. Methods 207:61-67, 1997).

5. Validation of the targets by protein-chemical methods using antibodies

The antibodies which can be produced as described above can be used to make a number of important statements about the target protein. Specifically the following analyses of validating the target protein are useful:

Specificity of the antibody
Assays based on cell culture with subsequent Western blotting are most suitable for demonstrating the fact that an antibody binds specifically only to the desired target protein (various variations are described, for example, in "Current Protocols in Proteinchemistry", John Wiley & Sons Ltd., Wiley InterScience). For the demonstration, cells are transfected with a cDNA for the target protein, which is under the control of a strong eukaryotic promoter (e.g. cytomegalovirus promoter; CMV). A wide variety of methods (e.g. electroporation, liposome-based transfection, calcium phosphate precipitation) are well established for transfecting cell lines with DNA (e.g. Lemoine et al., *Methods Mol. Biol.* 75: 441-7, 1997). As an alternative, it is also possible to use cell lines which express the target gene endogenously (detection via target gene-specific RT-PCR). As a control, in the ideal case, homologous genes are cotransfected in the experiment, in order to be able to demonstrate in the following Western blot the specificity of the analyzed antibody.

In the subsequent Western blotting, cells from cell culture or tissue samples which might contain the target protein are lysed in a 1% strength SDS solution,

and the proteins are denatured in the process. The lysates are fractionated according to size by electrophoresis on 8-15% strength denaturing polyacrylamide gels (contain 1% SDS) (SDS polyacrylamide gel electrophoresis, SDS-PAGE). The proteins are then transferred by one of a plurality of blotting methods (e.g. semi-dry electroblot; Biorad) to a specific membrane (e.g. nitrocellulose, Schleicher & Schüll). The desired protein can be visualized on this membrane. For this purpose, the membrane is first incubated with the antibody which recognizes the target protein (dilution approx. 1:20-1:200, depending on the specificity of said antibody), for 60 minutes. After a washing step, the membrane is incubated with a second antibody which is coupled to a marker (e.g. enzymes such as peroxidase or alkaline phosphatase) and which recognizes the first antibody. It is then possible to make the target protein visible on the membrane in a color or chemiluminescent reaction (e.g. ECL, Amersham Bioscience). An antibody with a high specificity for the target protein should in the ideal case only recognise the desired protein itself.

Localization of the target protein

Various methods are used to confirm the membrane localization, identified in the in silico approach, of the target protein. An important and well-established method using the antibodies described above is immuno-fluorescence (IF). For this purpose, cells of established cell lines which either synthesize the target protein (detection of the RNA by RT-PCR or of the protein by Western blotting) or else have been transfected with plasmid DNA are utilized. A wide variety of methods (e.g. electroporation, liposome-based transfection, calcium phosphate precipitation) are well established for transfection of cell lines with DNA (e.g. Lemoine et al., *Methods Mol. Biol.* 75: 441-7, 1997). The plasmid transfected into the cells, in immunofluorescence, may encode the unmodified

protein or else couple different amino acid markers to the target protein. The principle markers are, for example, the fluorescent green fluorescent protein (GFP) in various differentially fluorescent forms, short peptide sequences of 6-12 amino acids for which high-affinity and specific antibodies are available, or the short amino acid sequence Cys-Cys-X-X-Cys-Cys which can bind via its cysteines specific fluorescent substances (Invitrogen). Cells which synthesize the target protein are fixed, for example, with paraformaldehyde or methanol. The cells may then, if required, be permeabilized by incubation with detergents (e.g. 0.2% Triton X-100). The cells are then incubated with a primary antibody which is directed against the target protein or against one of the coupled markers. After a washing step, the mixture is incubated with a second antibody coupled to a fluorescent marker (e.g. fluorescein, Texas Red, Dako), which binds to the first antibody. The cells labeled in this way are then overlaid with glycerol and analyzed with the aid of a fluorescence microscope according to the manufacturer's information. Specific fluorescence emissions are achieved in this case by specific excitation depending on the substances employed. The analysis usually permits reliable localization of the target protein, the antibody quality and the target protein being confirmed in double stainings with, in addition to the target protein, also the coupled amino acid markers or other marker proteins whose localization has already been described in the literature being stained. GFP and its derivatives represent a special case, being excitable directly and themselves fluorescing. The membrane permeability which may be controlled through the use of detergents, in immunofluorescence, allows demonstration of whether an immunogenic epitope is located inside or outside the cell. The prediction of the selected proteins can thus be supported experimentally. An alternative possibility is to detect extracellular domains by means of flow

cytometry. For this purpose, cells are fixed under non-permeabilizing conditions (e.g. with PBS/Na azide/2% FCS/5 mM EDTA) and analyzed in a flow cytometer in accordance with the manufacturer's instructions. Only extracellular epitopes can be recognized by the antibody to be analyzed in this method. A difference from immunofluorescence is that it is possible to distinguish between dead and living cells by using, for example, propidium iodide or Trypan blue, and thus avoid false-positive results.

Another important detection is by immunohistochemistry (IHC) on specific tissue samples. The aim of this method is to identify the localization of a protein in a functionally intact tissue aggregate. IHC serves specifically for (1) being able to estimate the amount of target protein in tumor and normal tissues, (2) analyzing how many cells in tumor and healthy tissues synthesize the target gene, and (3) defining the cell type in a tissue (tumor, healthy cells) in which the target protein is detectable. Alternatively, the amounts of protein of a target gene may be quantified by tissue immunofluorescence using a digital camera and suitable software (e.g. Tillvision, Till-photonics, Germany). The technology has frequently been published, and details of staining and microscopy can therefore be found, for example, in "Diagnostic Immunohistochemistry" by David J., MD Dabbs ISBN: 0443065667 or in "Microscopy, Immunohistochemistry, and Antigen Retrieval Methods: For Light and Electron Microscopy" ISBN: 0306467704. It should be noted that, owing to the properties of antibodies, different protocols have to be used (an example is described below) in order to obtain a meaningful result.

Normally, histologically defined tumor tissues and, as reference, comparable healthy tissues are employed in IHC. It is also possible to use as positive and negative controls cell lines in which the presence of

the target gene is known through RT-PCR analyses. A background control must always be included.

Formalin-fixed (another fixation method, for example with methanol, is also possible) and paraffin-embedded tissue pieces with a thickness of 4 µm are applied to a glass support and deparaffinated with xylene, for example. The samples are washed with TBS-T and blocked in serum. This is followed by incubation with the first antibody (dilution: 1:2 to 1:2000) for 1-18 hours, with affinity-purified antibodies normally being used. A washing step is followed by incubation with a second antibody which is coupled to an alkaline phosphatase (alternative: for example peroxidase) and directed against the first antibody, for approx. 30-60 minutes. This is followed by a color reaction using said alkaline phosphatase (cf., for example, Shi et al., *J. Histochem. Cytochem.* 39: 741-748, 1991; Shin et al., *Lab. Invest.* 64: 693-702, 1991). To demonstrate antibody specificity, the reaction can be blocked by previous addition of the immunogen.

Analysis of protein modifications

Secondary protein modifications such as, for example, N- and O-glycosylations or myristilations may impair or even completely prevent the accessibility of immunogenic epitopes and thus call into question the efficacy of antibody therapies. Moreover, it has frequently been demonstrated that the type and amount of secondary modifications differ in normal and tumor tissues (e.g. Durand & Seta, 2000; Clin. Chem. 46: 795-805; Hakomori, 1996; Cancer Res. 56: 5309-18). The analysis of these modifications is therefore essential to the therapeutic success of an antibody. Potential binding sites can be predicted by specific algorithms.

Analysis of protein modifications usually takes place by Western blotting (see above). Glycosylations which usually have a size of several kDa, especially lead to

a larger total mass of the target protein, which can be fractionated in SDS-PAGE. To detect specific O- and N-glycosidic bonds, protein lysates are incubated prior to denaturation by SDS with O- or N-glycosylases (in accordance with their respective manufacturer's instructions, e.g. PNgase, endoglycosidase F, endoglycosidase H, Roche Diagnostics). This is followed by Western blotting as described above. Thus, if there is a reduction in the size of a target protein after incubation with a glycosidase, it is possible to detect a specific glycosylation and, in this way, also analyze the tumor specificity of a modification.

Functional analysis of the target gene

The function of the target molecule may be crucial for its therapeutic usefulness, so that functional analyses are an important component in the characterization of therapeutically utilizable molecules. The functional analysis may take place either in cells in cell culture experiments or else in vivo with the aid of animal models. This involves either switching off the gene of the target molecule by mutation (knockout) or inserting the target sequence into the cell or the organism (knockin). Thus it is possible to analyze functional modifications in a cellular context firstly by way of the loss of function of the gene to be analyzed (loss of function). In the second case, modifications caused by addition of the analyzed gene can be analyzed (gain of function).

a. Functional analysis in cells

Transfection. In order to analyze the gain of function, the gene of the target molecule must be transferred into the cell. For this purpose, cells which allow synthesis of the target molecule are transfected with a DNA. Normally, the gene of the target molecule here is under the control of a strong eukaryotic promoter (e.g. cytomegalovirus promoter; CMV). A wide variety of methods (e.g. electroporation, liposome-based

transfection, calcium phosphate precipitation) are well established for transfecting cell lines with DNA (e.g. Lemoine et al., *Methods Mol. Biol.* 75: 441-7, 1997). The gene may be synthesized either transiently, without genomic integration, or else stably, with genomic integration after selection with neomycin, for example.

RNA interference (siRNA). An inhibition of expression of the target gene, which may induce a complete loss of function of the target molecule in cells, may be generated by the RNA interference (siRNA) technology in cells (Hannon, GJ. 2002. RNA interference. Nature 418: 244-51; Czauderna et al. 2003. Nucl. Acid Res. 31: 670-82). For this purpose, cells are transfected with short, double-stranded RNA molecules of approx. 20-25 nucleotides in length, which are specific for the target molecule. An enzymic process then results in degradation of the specific RNA of the target gene and thus in an inhibition of the function of the target protein and consequently enables the target gene to be analyzed.

Cell lines which have been modified by means of transfection or siRNA may subsequently be analyzed in different ways. The most common examples are listed below.

1. Proliferation and cell cycle behavior

A multiplicity of methods for analyzing cell proliferation are established and are commercially supplied by various companies (e.g. Roche Diagnostics, Invitrogen; details of the assay methods are described in the numerous application protocols). The number of cells in cell culture experiments can be determined by simple counting or by colorimetric assays which measure the metabolic activity of the cells (e.g. wst-1, Roche Diagnostics). Metabolic assay methods measure the number of cells in an experiment indirectly via enzymic markers. Cell proliferation may be measured directly by analyzing

the rate of DNA synthesis, for example by adding bromodeoxyuridine (BrdU), with the integrated BrdU being detected colorimetrically via specific antibodies.

2. Apoptosis and cytotoxicity

A large number of assay systems for detecting cellular apoptosis and cytotoxicity are available. A decisive characteristic is the specific, enzyme-dependent fragmentation of genomic DNA, which is irreversible and results in certain death of the cell. Methods for detecting these specific DNA fragments are commercially obtainable. An additional method available is the TUNEL assay which can detect DNA single-strand breaks also in tissue sections. Cytotoxicity is mainly detected via an altered cell permeability which serves as marker of the vitality state of cells. This involves on the one hand the analysis of markers which can typically be found intracellularly in the cell culture supernatant. On the other hand, it is also possible to analyze the absorbability of dye markers which are not absorbed by intact cells. The best-known examples of dye markers are Trypan blue and propidium iodide, a common intracellular marker is lactate dehydrogenase which can be detected enzymatically in the supernatant. Different assay systems of various commercial suppliers (e.g. Roche Diagnostics, Invitrogen) are available.

3. Migration assay

The ability of cells to migrate is analyzed in a specific migration assay, preferably with the aid of a Boyden chamber (Corning Costar) (Cinamon G., Alon R. J. Immunol. Methods. 2003 Feb; 273(1-2):53-62; Stockton et al. 2001. Mol. Biol. Cell. 12: 1937-56). For this purpose, cells are cultured on a filter with a specific pore size. Cells which can migrate are capable of migrating through this filter into another culture vessel below. Subsequent microscopic analysis then permits determination of a possibly altered migration

behavior induced by the gain of function or loss of function of the target molecule.

b. Functional analysis in animal models

A possible alternative of cell culture experiments for the analysis of target gene function are complicated in vivo experiments in animal models. Compared to the cell-based methods, these models have the advantage of being able to detect faulty developments or diseases which are detectable only in the context of the whole organism. A multiplicity of models for human disorders are available by now (Abate-Shen & Shen. 2002. Trends in Genetics S1-5; Matsusue et al. 2003. J. Clin. Invest. 111:737-47). Various animal models such as, for example, yeast, nematodes or zebra fish have since been characterized intensively. However, models which are preferred over other species are mammalian animal models such as, for example, mice (Mus musculus) because they offer the best possibility of reproducing the biological processes in a human context. For mice, on the one hand transgenic methods which integrate new genes into the mouse genome have been established in recent years (gain of function; Jegstrup I. et al. 2003. Lab Anim. 2003 Jan.; 37(1):1-9). On the other hand, other methodical approaches switch off genes in the mouse genome and thus induce a loss of function of a desired gene (knockout models, loss of function; Zambrowicz BP & Sands AT. 2003. Nat. Rev. Drug Discov. 2003 Jan; 2(1):38-51; Niwa H. 2001. Cell Struct. Funct. 2001 Jun; 26(3):137-48); technical details have been published in large numbers.

After the mouse models have been generated, alterations induced by the transgene or by the loss of function of a gene can be analyzed in the context of the whole organism (Balling R, 2001. Ann. Rev. Genomics Hum. Genet. 2:463-92). Thus it is possible to carry out, for example, behavior tests as well as to biochemically study established blood parameters. Histological

analyses, immunohistochemistry or electron microscopy enable alterations to be characterized at the cellular level. The specific expression pattern of a gene can be detected by in-situ hybridization (Peters T. et al. 2003. Hum. Mol. Genet 12:2109-20).

Example 3: Identification of SEQ ID NO: 1/2 as a diagnostic and therapeutic cancer target

SEQ ID NO: 1 (nucleic acid sequence) is encoded by a new gene on chromosome 6 (6q26-27) and represents the deduced protein sequence (SEQ ID NO: 2). An alternative open reading frame of this gene locus is SEQ ID NO: 267 which codes for the deduced protein sequence SEQ ID NO: 268. Both protein sequences show no homologies to previously known proteins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing a specific quantitative RT-PCR (primer pair SEQ ID NO: 3 and 4). The transcript was not detected in any of the normal tissues analyzed. Surprisingly, we detected very specifically substantial amounts of said transcript in almost all melanoma samples studied, although the gene is not expressed in normal skin as tissue of origin (fig. 1). The selectivity of this marker for melanomas was confirmed by a conventional RT-PCR (fig. 2). Surprisingly, we amplified in the process two fragments which reflect gene-specific variants (probably SEQ ID NO: 1 and SEQ ID NO: 267).

We thus demonstrate that this gene is an absolutely specific marker for melanoma cells and, due to its absence in each of the normal tissues studied, is suitable as biomarker for targeted therapeutic and diagnostic approaches.

In particular it is possible to utilize according to the invention extracellular portions of SEQ ID NO: 2 or 268 as target structure of monoclonal antibodies. This applies inter alia to the following epitopes: amino acids 1-50 based on SEQ ID NO: 2, amino acids 1-12

based on SEQ ID NO: 268, amino acids 70-88 based on SEQ ID NO: 2, amino acids 33-129 based on SEQ ID NO: 268, and SEQ ID NO: 281.

According to the invention, other target-oriented approaches such as vaccines and therapies with small compounds, which have only this gene as target structure and thus do not affect any healthy cells, are also therapeutically conceivable. Said gene may also be utilized diagnostically owing to its selectivity for tumor cells.

Example 4: Identification of SEQ ID NO: 5/6 as diagnostic and therapeutic cancer target

SEQ ID NO: 5 (nucleic acid sequence) is encoded by a new gene on chromosome 11 (11q12.1) and represents the deduced protein sequence (SEQ ID NO: 6). An alternative open reading frame of this gene locus is SEQ ID NO: 269 which codes for the deduced protein sequence SEQ ID NO: 270. Both protein sequences show no homologies to previously known proteins.

According to the invention, the amount of gene-specific transcript in healthy tissue and in carcinoma samples (in each case pool of samples) was studied after establishing a gene-specific quantitative RT-PCR (primer pair SEQ ID NO: 7 and 8). We detected no specific RNA at all or else only small amounts thereof in the healthy tissues we studied, with the exception of testis (fig. 3; A quantitative RT-PCR; B gel image). Consequently, there is a high probability of the locus expressing a germ cell-specific gene product. However, the gene is activated in many tumor samples, and specific RNA was detectable in substantial amounts (fig. 3). The highest prevalence and level of expression were found in renal cell tumors. But specific transcripts were also detectable in gastric, pancreatic, ENT and lung tumors (fig. 4; A quantitative RT-PCR; B gel image). Even repeated examinations of the corresponding normal tissues were unable to detect gene-specific transcripts. In order to additionally

prove expression from this gene locus, a Northern blot was additionally carried out. For this purpose, a probe was prepared in a specific PCR of primers SEQ ID NO: 7 and 8 with incorporation of digoxigenin-dUTP (Roche Diagnostics) according to the manufacturer's instructions. The probe was then hybridized with 2 µg (fig. 5, lane 1) and 1 µg (fig. 5, lane 2), respectively, of total RNA from testis tissue and the digoxigenin of said probe was subsequently detected in a specific color reaction. An approx. 3.1 kB gene-specific fragment was detected in the experiment (fig. 5) and thus additionally confirmed expression of this locus.

Said gene locus is thus a typical representative of the class of the "cancer/testis antigens" which are expressed in normal tissues virtually exclusively in the germ cells of the testis. In tumors, however, cancer/testis antigens are frequently switched on, although they are not expressed in the underlying somatic normal tissue cells. Several members of this functionally and structurally heterogeneous class are already tested for specific immunotherapeutic approaches with cancers in phase I/II studies, owing to their attractive selective tissue distribution (e.g. Scanlan MJ, Gure AO, Jungbluth AA, Old LJ, Chen YT. 2002. Immunol. Rev. 2002 Oct; 188:22-32).

Antibodies may be produced by utilizing the peptides according to SEQ ID NO: 282 and 283. In particular, according to the invention it is possible to utilize the extracellular domains of SEQ ID NO: 6 and SEQ ID NO: 270 as target structures of monoclonal antibodies.

Example 5: Identification of LOC203413 as diagnostic and therapeutic cancer target

The gene or protein of the gene locus LOC203413 (nucleic acid sequence: SEQ ID NO: 9; amino acid sequence: SEQ ID NO: 10) is a gene on the X chromosome (Xq24), which has not been characterized previously. Aside from a transmembrane domain, it has no further

functional motifs and no homologies to previously known proteins.

According to the invention, the amount of transcript in healthy tissue and in carcinoma samples (pool of samples, number indicated in the figure) was studied after establishing an LOC203413-specific quantitative RT-PCR (primer pair SEQ ID NO: 11 and 12) (fig. 6; A: quantitative evaluation, B: image after gel-electrophoretic fractionation). LOC203413-specific RNA cannot be detected in any of the healthy tissues we studied, with the exception of testis. Consequently, it is highly probable that LOC203413 is a germ cell-specific gene product. As fig. 6 reveals, LOC203413-specific transcripts were detectable in gastric, pancreatic, esophageal, mammary, ovarian and prostate carcinomas, with high expression being observed in particular in gastric and mammary carcinomas. For a more detailed analysis, healthy gastric samples and gastric carcinoma samples were additionally characterized in a quantitative RT-PCR (fig. 7A). LOC203413 was expressed in 70% of the carcinomas, whereas no significant expression was detectable in any of the healthy gastric samples. The MKN45 cell line which is derived from a gastric carcinoma also expresses LOC203413. In addition, specific expression was detected in 2/3rds of pancreatic tumors studied and in 40% of liver carcinomas (fig. 7B).

LOC203413 is thus a typical representative of the class of cancer/testis antigens which are expressed in normal tissues exclusively in the germ cells of the testis. In tumors, however, cancer/testis antigens are frequently switched on, although they are not expressed in the underlying somatic normal tissue cells. Several members of this functionally and structurally heterogeneous class are already tested for specific immunotherapeutic approaches with cancers in phase I/II studies, owing to their attractive selective tissue distribution (e.g. Scanlan MJ, Gure AO, Jungbluth AA, Old LJ, Chen YT. 2002. Immunol. Rev. 2002 Oct; 188:22-32).

In particular it is possible to utilize according to the invention the extracellular domain of LOC203413 as target structure of monoclonal antibodies. Thus the amino acids 22-113 (SEQ ID NO: 284) are of interest as epitopes. Conserved N-glycosylation motifs are located in the sequence at amino acid positions 34 and 83, based on SEQ ID NO: 10, which motifs may be suitable in particular for producing tumor-specific antibodies. LOC203413-specific antibodies were produced by using the peptides listed under SEQ ID NO: 285 and 286.
According to the invention, other target-oriented approaches such as vaccines and therapies with small compounds, which have only this gene as target structure and thus do not affect any healthy cells, are also therapeutically conceivable. Said gene may also be utilized diagnostically owing to its selectivity for tumor cells.

Example 6: Identification of LOC90625 as a diagnostic and therapeutic cancer target

The gene LOC90625 (nucleic acid sequence: SEQ ID NO: 13) is a gene on chromosome 21 (21q22.3), which has not been characterized previously. It encodes a protein (amino acid sequence: SEQ ID NO: 14) having a trans-membrane domain but otherwise no homologies to previously known proteins.
According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples (pool of samples, the number is indicated in the figure) was investigated after establishing an LOC90625-specific quantitative RT-PCR (primer pair SEQ ID NO: 15 and 16) (fig. 8). LOC90625 is expressed very selectively in healthy tissue, with specific transcripts being detectable especially in testis. In all other healthy tissues analyzed LOC90625-specific expression was detectable only at a low level, if at all (fig. 8). Surprisingly, we detected LOC90625-specific overexpression in some types of tumors.

LOC90625 was strongly overexpressed in particular in prostate, esophageal and pancreatic carcinomas, in comparison to the respective healthy tissue samples (figs. 8 and 9A).

LOC90625 is a selectively expressed antigen which is obviously increasingly expressed in proliferating tissues. Thus a selective overexpression in tumors can be observed which is therapeutically utilizable.

The extracellular domain of LOC90625 in particular can be utilized according to the invention as target structure of monoclonal antibodies. Said structure may be, for example, 1-19 (SEQ ID NO: 287) or else the amino acids 40-160 (SEQ ID NO: 288). LOC203413-specific antibodies were produced by using the peptides according to SEQ ID NO: 289 and 290.

Example 7: Identification of the FAM26A protein as a diagnostic and therapeutic cancer target

The FAM26A gene (SEQ ID NO: 17; NM_182494) which is located on chromosome 10 (10q24) encodes the gene product of SEQ ID NO: 18 (NP_872300). FAM26A has several transmembrane domains, with an N-glycosylation motif being located at amino acid position 142. The deduced protein sequence displays a distant homology to the PMP/claudin family.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in tumor samples was investigated after establishing an FAM26A-specific quantitative RT-PCR (primer pair SEQ ID NO: 19 and 20) (fig. 10). Surprisingly, we were able to detect overexpression of FAM26A in various tumors. FAM26A was expressed at a distinctly higher level in particular in ovarian, gastric, esophageal, pancreatic and liver tumors, in comparison with the corresponding healthy tissue. According to the invention, selectively high expression of FAM26A in various tumor tissues may be utilized for molecular diagnostic methods such as, for example, RT-PCR for detecting tumor cells in tissue biopsies.

In order to further verify the expression data, FAM26A-specific antibodies were produced by immunization of animals. Polyclonal antibodies were produced by using the peptides listed under SEQ ID NO: 291 and 292. The specificity of the antibodies was demonstrated by Western blot analysis (fig. 11A: SEQ ID NO: 291; B: SEQ ID NO: 292). For this purpose, COS cells were transfected with an FAM26 fragment-encoding plasmid construct. The Western blot showed a specific signal with both antibodies, which was not detectable in the respective controls (fig. 11). We detected FAM26A also in various cervical, ovarian and pancreatic tumors, using a SEQ ID NO: 292-specific antibody (fig. 12), as well as in the cell lines SW480, EFO 27 and SNU 16 which were in each case RT-PCR-positive, using a SEQ ID NO: 291-specific antibody (fig. 13A). Here we found, in addition to an approx. 50 kDa specific band, also a weaker band at approx. 40 kDa. The latter corresponds to about the expected size. The major fragment at 50 kDa represents a post-translationally modified protein. The endogenous FAM26A protein was moreover detected in SW480 cells by means of immunofluorescence using a SEQ ID NO: 292-specific antibody. The analysis reveals localization in the plasma membrane (fig. 13B). In order to analyze localization of FAM26A in a tissue assemblage, healthy testis samples were characterized immunohistologically. In testis, the FAM26A protein was detected specifically in the membrane of spermatocytes, and due to the results, a membrane localization of FAM26A appears likely (fig. 14). This was also confirmed in tumor samples (fig. 15).

The extracellular domains of FAM26A in particular may be utilized according to the invention as target structures of monoclonal antibodies. These are, based on SEQ ID NO: 17, the amino acids 38-48 (SEQ ID NO: 293) and the amino acids 129-181 (SEQ ID NO: 294). Alternatively, the C-terminal amino acids 199-334 (SEQ ID NO: 295) may also be preferred epitopes for

producing antibodies for diagnostic or therapeutic purposes. In addition, the N-glycosylation motif at position 142 may be an interesting point of attack for therapeutic antibodies.

Example 8: Identification of SEMA5B as diagnostic and therapeutic cancer target

The gene semaphorin 5B (SEMA5B; SEQ ID NO: 21) which encodes the protein of SEQ ID NO: 22 is located on chromosome 3 (3q21.1). SEMA5B is a type I transmembrane protein and belongs to the family of semaphorins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples (pool of samples, the number is indicated in the figure) was investigated after establishing an SEMA5B-specific quantitative RT-PCR (primer pair SEQ ID NO: 23 and 24) (fig. 16). We found that, in healthy tissue, SEMA5B is very selectively restricted to testis and skin. In all other healthy tissues analyzed SEMA5B-specific expression was detectable at low level or not at all (fig. 16). In contrast, we surprisingly found SEMA5B-specific overexpression in some types of tumors, in particular in kidney carcinomas and breast tumors (fig. 17A and B), in comparison to the respective healthy tissues.

Said selective overexpression in tumors can be utilized therapeutically.

The extracellular domain of SEMA5B (aa 20-1035; SEQ ID NO: 296) in particular may be utilized according to the invention as target structure of antibodies. SEMA5B is a type I transmembrane domain protein (TM aa 1035-1057) whose C terminus is located inside the cell (aa 1058-1151). SEMA5B-specific antibodies were produced by using the peptides according to SEQ ID NO: 297 and 298.

Example 9: Identification of GJB5 as a diagnostic and therapeutic cancer target

The protein GBJ5 (nucleic acid sequence: SEQ ID NO: 25; amino acid sequence: SEQ ID NO: 26) is a member of the

connexin family. The gene consists of two exons and is located on chromosome 1 (1p35.1). The deduced amino acid sequence codes for a protein of 273 amino acids. Connexins have an important function in cell-cell contacts via "gap junctions" which are used for exchanging small cytoplasmic molecules, ions and secondary transmitters and thus enable individual cells to communicate with each other. Gap junctions consist of several connexin subunits which form a membrane channel. 11 different members of the connexins have been described to date, all of which are located in a gene cluster on chromosome 1 (Richard, G.; Nature Genet. 20: 366-369, 1998). GBJ5 has four transmembrane domains, with the N and C termini of the protein being located inside the cell.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated (pool of samples, the number is indicated in the figure) after establishing a GBJ5-specific quantitative RT-PCR (primer pair SEQ ID NO: 27, 28). Our studies reveal differential distribution of expression in normal tissues. We found GBJ5 transcripts to be expressed virtually exclusively in the esophagus and in the skin, with transcription being very weak or not detectable in all other tissues analyzed (fig. 18). Very strong tumor-specific overexpression was observed in esophageal, colon, gastric and pancreatic carcinomas (fig. 18). This was confirmed by analyzing individual samples of the four carcinomas (fig. 19 A-D). In addition, the GBJ5-specific transcript can clearly be detected in the established cell lines LoVo, MKN45 and NCI-N87 (fig. 19 A-D).

The extracellular domains of GBJ5 in particular may be utilized according to the invention as target structure of therapeutic antibodies. Based on SEQ ID NO: 26, the amino acids 41-75 (SEQ ID NO: 299) and the region between amino acids 150 and 187 (SEQ ID NO: 300) are located extracellularly. GBJ5-specific antibodies were

produced by using the peptides according to SEQ ID NO: 301 and 302.

Example 10: Identification of KLK5 as a diagnostic and therapeutic cancer target

The gene KLK5 (SEQ ID NO: 29) and its translation product (SEQ ID NO: 30) is a member of the kallikrein family, a group of serine proteases with very different physiological functions. The gene is located on chromosome 19 (19q13.3-13.4) and codes for a serine protease. KLK5 is synthesized as pro form and is activated by proteolysis in the stratum corneum (Brattsand, M et al; *J. Biol. Chem.* 274: 1999). The active protease (aa 67-293) is secreted and is involved in the process of desquamation. The propeptide (aa 30-67) remains bound to the cell surface via the transmembrane domain (aa 1-29) (Ekholm, E et al; *Jour Investigative Dermatol*, 114; 2000).

According to the invention the distribution of KLK5-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing a KLK5-specific quantitative RT-PCR (primer pair SEQ ID NO: 31, 32) (fig. 20). In most normal tissues expression of KLK5 is at a very low to non-existent level, with moderate expression of KLK5 being found only in testis, esophagus, skin and prostate. We detected significant overexpression of KLK5 in esophageal carcinomas, cervical and in ENT tumors, in comparison with the corresponding normal tissues of origin (fig. 20, 21). Distinctly weaker but detectable KLK5-specific expression was moreover detected in some tumors of other tissues (e.g. in gastric and pancreatic carcinomas).

The extracellular domain of KLK5 in particular may be utilized according to the invention as target structure of therapeutic antibodies (SEQ ID NO: 303). The region of the propeptide (amino acids 30 to 67) is particularly suitable for this. KLK5-specific anti-

bodies were produced by using the peptide listed under SEQ ID NO: 304.

Example 11: Identification of LOC352765 as a diagnostic and therapeutic cancer target

The LOC352765 gene locus is located on chromosome 9 (9q34.12). The gene (SEQ ID NO: 33) encodes the gene product of SEQ ID NO: 34. The LOC352765 protein has a transmembrane domain at the N terminus. The hypothetical protein displays no homologies to previously known proteins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples (pool of samples) was investigated after establishing an LOC352765-specific quantitative RT-PCR (primer pair SEQ ID NO: 35 and 36) (fig. 22). LOC352765 is expressed very selectively in healthy tissue, and we found specific transcripts to be detectable only in testis, skin and bladder. In contrast, LOC352765-specific overexpression was detected in some types of tumors. Particularly in breast tumors, expression was higher than in the normal tissue with the highest level of expression. We also found LOC352765 to be distinctly overexpressed in colon and ovarian carcinomas and in ENT tumors (figs. 22, 23).

Owing to its selective overexpression in tumors, LOC352765 can be utilized therapeutically. The extracellular domain of LOC352765 (amino acids 44-211, SEQ ID NO: 34) in particular may be utilized according to the invention as target structure of antibodies and other targeted forms of therapy. Specific antibodies were produced by using the peptides according to SEQ ID NO: 305 and 306.

Example 12: Identification of SVCT1 as a diagnostic and therapeutic cancer target

The gene SVCT1 (SEQ ID NO: 37) is located on chromosome 7 (7q33) and codes for the gene product of SEQ ID NO: 38. The SVCT1 protein has four transmembrane

domains and displays no homologies to previously known proteins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples (pool of samples) was investigated after establishing an SVCT1-specific quantitative RT-PCR (primer pair SEQ ID NO: 39 and 40) (fig. 24). SVCT1 in healthy tissue is restricted selectively to kidney, testis, thymus and mammary gland. In contrast, SVCT1-specific overexpression was surprisingly detected in some types of tumors. SVCT1 is strongly overexpressed in particular in carcinomas of the kidney, esophagus and pancreas and in ENT tumors (figs. 24, 25), and that is not only in comparison with the corresponding healthy tissue of origin but also with respect to the normal tissue with the highest level of expression over all.

SVCT1 can be therapeutically utilized owing to its selective overexpression in tumors. The extracellular domains of SVCT1 in particular may be utilized according to the invention as target structures of antibodies and for other targeted forms of therapy. Specific antibodies were produced by using the peptides according to SEQ ID NO: 307 and 308.

Example 13: Identification of LOC199953 as a diagnostic and therapeutic cancer target

The gene or protein of the LOC199953 gene locus (nucleic acid sequence: SEQ ID NO: 41; amino acid sequence: SEQ ID NO: 42) is located on chromosome 1 (1q36.22). The protein has several transmembrane domains. Alternative open reading frames of this gene locus are SEQ ID NO: 271 with its gene product SEQ ID NO: 272 and SEQ ID NO: 273 with the corresponding gene product SEQ ID NO: 274. Other than that, the hypothetical protein displays no further homologies to previously known protein domains.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing an LOC199953-

specific quantitative RT-PCR (primer pair SEQ ID NO: 43 and 44). LOC199953 is selectively expressed in healthy tissues and overexpressed in some tumors. In particular, it was possible to identify overexpression in ENT and kidney carcinomas (fig. 26) in approx. 50% of the tumor samples, in comparison with normal tissues.

According to the invention, the extracellular domains of LOC199953 may be utilized as target structure of antibodies.

Example 14: Identification of TMEM31 as a diagnostic and therapeutic cancer target

The gene TMEM31 (SEQ ID NO: 45) of the LOC203562 gene locus is located on chromosome X (Xq22.2). The gene codes for the protein of SEQ ID NO: 46. Said protein has two transmembrane domains and otherwise displays no homologies to previously known proteins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing a TMEM31-specific quantitative RT-PCR (primer pair SEQ ID NO: 47 and 48). In healthy tissues, TMEM31 is very selectively restricted especially to testis (fig. 27). Surprisingly, we also found expression in some types of tumors, whereas no expression was detectable in the corresponding normal tissues. Said tumors are in particular carcinomas of the kidney, colon, stomach, breast, liver and lung and ENT carcinomas (figs. 27, 28).

TMEM31 is thus a typical representative of the class of cancer/testis antigens which are expressed in normal tissues exclusively in the germ cells of the testis. In tumors, however, cancer/testis antigens are frequently switched on, although they are not expressed in the underlying somatic normal tissue cells. Several members of this functionally and structurally heterogeneous class are already tested for specific immunotherapeutic approaches with cancers in phase I/II studies, owing to

their attractive selective tissue distribution (e.g. Scanlan MJ, Gure AO, Jungbluth AA, Old LJ, Chen YT. 2002. Immunol. Rev. 2002 Oct; 188:22-32).

The extracellular TMEM31 domains may be utilized according to the invention as target structure of antibodies.

Example 15: Identification of FLJ25132 as a diagnostic and therapeutic cancer target

The FLJ25132 gene/protein (nucleic acid sequence: SEQ ID NO: 49; amino acid sequence: SEQ ID NO: 50) is located on chromosome 17 (17q25.3). FLJ25132 has a transmembrane domain but otherwise does not display any homologies to previously known proteins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing an FLJ25132-specific quantitative RT-PCR (primer pair SEQ ID NO: 51 and 52). FLJ25132 is partially overexpressed in the carcinoma samples studied by us, in comparison to healthy tissue (fig. 29). Distinct overexpression of FLJ25132 was detected in particular in ovarian and in prostate carcinomas.

The extracellular FLJ25132 domains may be utilized according to the invention as target structure of antibodies.

Example 16: Identification of LOC143724, LOC284263, LOC283435 and LOC349260 as diagnostic and therapeutic cancer targets

The gene loci (with the correspondingly encoded genes and gene products), LOC143724, LOC284263, LOC283435 and LOC349260, are combined, owing to their similar profiles.

The gene with SEQ ID NO: 53, which is present in the LOC143724 gene locus on chromosome 11 (11q13.1), encodes the gene product SEQ ID NO: 54. SEQ ID NO: 275 with its gene product SEQ ID NO: 276 represents an alternative open reading frame of this gene locus,

which is either a separate transcript or a splice variant of SEQ ID NO: 53. The primers according to SEQ ID NO: 55 and 56 were used for gene-specific amplification of said gene.

The gene with SEQ ID NO: 89, which is present in the LOC284263 gene locus on chromosome 18 (18q21.1), encodes the gene product with SEQ ID NO: 90. The primers according to SEQ ID NO: 91 and 92 were used for gene-specific amplification of said gene.

The gene with SEQ ID NO: 117, which is present in the LOC283435 gene locus on chromosome 12 (12q24.32), encodes the gene product with SEQ ID NO: 118. The primers according to SEQ ID NO: 119 and 120 were used for gene-specific amplification of said gene.

The gene with SEQ ID NO: 121, which is present in the LOC349260 gene locus on chromosome 9 (9q11.2), encodes the gene product with SEQ ID NO: 122. The primers according to SEQ ID NO: 123 and 124 were used for gene-specific amplification of said gene.

All proteins have transmembrane domains and, in addition, do not display any homologies to previously known proteins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing specific quantitative RT-PCR analyses. None of the four genes were detected in the healthy tissues which are investigated, with the exception of testis. Consequently, there is a high probability of said genes being germ cell-specific. However, surprisingly significant expression is found in various tumor samples.

The four genes are thus typical representatives of the class of cancer/testis antigens which are expressed in normal tissues exclusively in the germ cells of the testis. In tumors, however, cancer/testis antigens are frequently switched on, although they are not expressed in the underlying somatic normal tissue cells. Several members of this functionally and structurally

heterogeneous class are already tested for specific immunotherapeutic approaches with cancers in phase I/II studies, owing to their attractive selective tissue distribution (e.g. Scanlan MJ, Gure AO, Jungbluth AA, Old LJ, Chen YT. 2002. Immunol. Rev. 2002 Oct; 188:22-32).

The extracellular domains of the four genes may be utilized according to the invention as target structure of antibodies.

Example 17: Identification of the sequence according to SEQ ID NO: 57 as a diagnostic and therapeutic cancer target

The sequence according to SEQ ID NO: 57 is derived from a gene on chromosome 1 (1p21.3) and encodes the protein sequence according to SEQ ID NO: 58. SEQ ID NO: 277 with its gene product SEQ ID NO: 278 represents an alternative transcript of said gene locus. The transmembrane protein does not display any homologies to previously known proteins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing a specific quantitative RT-PCR (primer pair SEQ ID NO: 59 and 60). SEQ ID NO: 57 is selectively expressed in the healthy tissues studied by us (fig. 30). Specific transcripts were detectable in nearly all types of tumors analyzed and overexpressed in particular in liver, ENT and kidney tumors. This was confirmed in the analysis of individual tumor samples in comparison with healthy tissue samples (fig. 31).

The extracellular domains of the sequence according to SEQ ID NO: 58 may be utilized according to the invention as target structure of antibodies, in particular with amino acids 20-38 and 90-133 being located extracellularly.

Example 18: Identification of LOC119395 as a diagnostic and therapeutic cancer target

The gene with SEQ ID NO: 61, which is present in the LOC119395 gene locus on chromosome 17 (17q25.3), encodes a gene product with SEQ ID NO: 62. The transmembrane protein displays no homologies to previously known proteins.

According to the invention the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing an LOC119395-specific quantitative RT-PCR (primer pair SEQ ID NO: 63 and 64) (fig. 32). LOC119395 is very selectively expressed in the healthy tissues studied by us and is detectable only in a few tissues (fig. 32). In contrast, LOC119395-specific transcripts were detectable in nearly all types of tumors analyzed. In parts distinct, tumor-selective overexpression of LOC119395 was observed in particular in gastric, ovarian and prostate carcinomas. This was confirmed in the analysis of individual tumor samples in comparison with healthy tissue samples (fig. 33). It was possible to detect overexpression of LOC119395 in mammary carcinomas and esophageal tumors in comparison with the respective healthy tissue. Tumor-selective expression was identified in colon carcinomas and gastric carcinomas (fig. 33).

The extracellular LOC119395 domain (amino acids 44-129) may be utilized according to the invention as target structure of antibodies.

Example 19: Identification of LOC121838 as a diagnostic and therapeutic cancer target

The gene which is located in the LOC121838 gene locus on chromosome 13 (13q14.11) and has the transcript of SEQ ID NO: 65 encodes the protein with SEQ ID NO: 66. The transmembrane protein displays no homologies to previously known proteins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing an LOC121838-specific quantitative RT-PCR (primer pair SEQ ID NO: 67

and 68) (fig. 34A). LOC121838 is very selectively expressed in the healthy tissues studied by us and is detectable only in a few tissues (fig. 34A and B). In contrast, LOC121838-specific transcripts were detectable in many types of tumors analyzed. We found distinct tumor-selective overexpression of LOC121838 in particular in ovarian and esophageal carcinomas.
The extracellular LOC121838 domains may be utilized according to the invention as target structure of antibodies.

Example 20: Identification of LOC221103 as a diagnostic and therapeutic cancer target

The gene which is localized in the LOC221103 gene locus on chromosome 11 (11q12.3) and has the transcript of SEQ ID NO: 69 encodes the protein with SEQ ID NO: 70. The transmembrane protein displays no homologies to previously known proteins.
According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing an LOC221103-specific quantitative RT-PCR (primer pair SEQ ID NO: 71 and 72). In the healthy tissues studied by us, LOC221103 is expressed only in the liver and otherwise not detectable (fig. 35). Surprisingly, LOC221103-specific transcripts are overexpressed in liver carcinomas (fig. 36).
The extracellular LOC221103 domains may be utilized according to the invention as target structure of antibodies.

Example 21: Identification of LOC338579 as a diagnostic and therapeutic cancer target

The gene which is localized in the LOC338579 gene locus on chromosome 10 (10q11.21) and has the transcript of SEQ ID NO: 73 encodes the protein with SEQ ID NO: 74. The transmembrane protein displays no homologies to previously known proteins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing an LOC338579-specific quantitative RT-PCR (primer pair SEQ ID NO: 75 and 76). We found expression in healthy tissues only in testis and, at a lower level, in the liver and the thymus. Surprisingly, we found LOC338579 overexpression in colon carcinomas and liver carcinomas in comparison with the healthy tissue (fig. 37).
The extracellular LOC338579 domains may be utilized according to the invention as target structure of antibodies.

Example 22: Identification of LOC90342 as a diagnostic and therapeutic cancer target

The gene which is located in the LOC90342 gene locus on chromosome 2 (2q11.2) and has the transcript of SEQ ID NO: 77 encodes the protein with SEQ ID NO: 78. The transmembrane protein includes a calcium-binding motif (CalB) which is conserved in protein kinase C and in various phospholipases.
According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing an LOC90342-specific quantitative RT-PCR (primer pair SEQ ID NO: 79 and 80) (fig. 38). We found LOC90342 only in a small number of healthy tissues, most of which are of little relevance with regard to toxicity (fig. 38). In contrast, we found LOC90342-specific transcripts in a multiplicity of the types of tumors analyzed. In parts distinctly tumor-selective overexpression of LOC90342 was observed in particular in gastric, liver, pancreatic, prostate, ovarian and lung carcinomas.
The membrane protein has a single transmembrane domain (aa 707-726). The extracellular LOC90342 domain may be utilized according to the invention as target structure of therapeutic antibodies.

Example 23: Identification of LRFN1 as a diagnostic and therapeutic cancer target

LRFN1 (SEQ ID NO: 81) is a gene which is localized on chromosome 19 (19q13.2). The gene codes for the protein of SEQ ID NO: 82. Said protein includes a transmembrane domain and displays homologies to the Myb DNA-binding domain and to a C2-type immunoglobulin domain.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing an LRFN1-specific quantitative RT-PCR (primer pair SEQ ID NO: 83 and 84). LRFN1 is very weakly expressed in most of the normal tissues studied, except for activated PBMC and brain (fig. 39). In contrast, we found LRFN1-specific transcripts to be increasingly detectable in some of the types of tumors analyzed. We found distinct tumor-selective overexpression of LRFN1 in particular in gastric, pancreatic, esophageal and mammary carcinomas, in comparison with the corresponding normal tissues.

The protein includes a transmembrane domain (aa 448-470). The extracellular LFRN1 domains may be utilized according to the invention as target structure of therapeutic antibodies.

Example 24: Identification of LOC285916 as a diagnostic and therapeutic cancer target

The gene which is localized in the LOC285916 gene locus on chromosome 7 (7p22.3) and has the transcript of SEQ ID NO: 85 encodes the protein with SEQ ID NO: 86. The transmembrane protein displays no homologies to previously known proteins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing an LOC285916-specific quantitative RT-PCR (primer pair SEQ ID NO: 87 and 88). In the healthy tissues studied by us, LOC285916 is expressed selectively in testis, with no or only little expression being detected by us in all other tissues studied (fig. 40A). Surprisingly, we

found LOC285916-specific transcripts in all types of tumors tested. Distinct tumor-specific overexpression was detectable in particular in mammary, esophageal, renal, ENT and lung carcinomas (figs. 40A and B). The extracellular LOC285916 domains (amino acids 42 to 93) may be utilized according to the invention as target structure of antibodies.

Example 25: Identification of MGC71744 as a diagnostic and therapeutic cancer target

The MGC71744 gene with SEQ ID NO: 93 on chromosome 17 (17p13.2) encodes the protein with SEQ ID NO: 94. The transmembrane protein displays no homologies to previously known proteins. According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples (pool of samples) was studied after establishing an MGC71744-specific quantitative RT-PCR (primer pair SEQ ID NO: 95 and 96) (fig. 41). MGC71744 is hardly expressed in healthy tissue. We found small amounts of specific transcripts only in the lung and in the spleen. The level of MGC71744-specific expression in all other healthy tissues analyzed was low or not detectable at all (fig. 41A). In contrast, we surprisingly found MGC71744-specific overexpression in some types of tumors, in particular in carcinomas of the kidney (figs. 41A & B), in comparison with healthy tissue. The extracellular domain of MGC71744 (N terminus, aa 67-85) in particular may be utilized according to the invention as target structure of antibodies.

Example 26: Identification of LOC342982 as a diagnostic and therapeutic cancer target

The gene which is localized in the LOC342982 gene locus on chromosome 19 (19p13.13) and has the transcript of SEQ ID NO: 97 encodes the protein with SEQ ID NO: 98. The transmembrane protein displays homologies to the carbohydrate binding domain of C-type lectins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples (pool of samples) was investigated after establishing an LOC342982-specific quantitative RT-PCR (primer pair SEQ ID NO: 99 and 100). LOC342982-specific RNA is selectively expressed, with only a low level of expression or no expression being detectable in many normal tissues analyzed (fig. 42). In contrast, nearly all of the classes of tumors tested exhibited overexpression which was partly tumor-specific. Primarily pancreatic, kidney, lung and mammary carcinomas exhibit very strong expression of LOC342982-specific RNA (fig. 42).

The extracellular domain of LOC342982 (amino acids 178-339) in particular may be utilized according to the invention as target structure of monoclonal antibodies.

Example 27: Identification of LOC343169/OR6F1 as a diagnostic and therapeutic cancer target

The gene OR6F1 which is localized in the LOC343169 gene locus on chromosome 1 (1q44) and has the transcript of SEQ ID NO: 101 encodes the protein with SEQ ID NO: 102. OR6F1 has several transmembrane domains and belongs to the family of olfactory receptors and thus to the large family of G protein-coupled receptors.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples (pool of samples) was investigated after establishing an LOC343169/OR6F1-specific quantitative RT-PCR (primer pair SEQ ID NO: 103 and 104) (fig. 43A). LOC343169/OR6F1 is very selectively expressed in healthy tissue, with specific transcripts being detectable especially in testis and spleen. The level of LOC343169/OR6F1-specific expression was low or not detectable at all in all other healthy tissues analyzed (fig. 43A). In contrast, LOC343169/OR6F1-specific overexpression was surprisingly detected is some types of tumors. Tumor-specific overexpression of LOC343169/OR6F1 is seen in particular in mammary,

ovarian, kidney, prostate, pancreatic and liver carcinomas (fig. 43A). An analysis of individual samples confirmed overexpression in ovarian carcinomas. LOC343169/OR6F1 is a selectively expressed antigen which is obviously increasingly expressed in proliferating tissues. Thus selective overexpression in tumors can be observed which is therapeutically utilizable. The extracellular domains in particular may be utilized according to the invention as target structures of monoclonal antibodies.

Example 28: Identification of LOC340204 as a diagnostic and therapeutic cancer target

The gene which is localized in the LOC340204 gene locus on chromosome 6 (6p21.31) and has the transcript of SEQ ID NO: 105 encodes the protein with SEQ ID NO: 106. Said protein has a transmembrane domain. Moreover said protein displays strong homology to a "colipase" domain. A cofactor function for pancreatic lipase is attributed to colipase. SEQ ID NO: 279 with its gene product SEQ ID NO: 280 represents an alternative transcript of said gene locus, which could be both a separate transcript and a splice variant of SEQ ID NO: 105.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing an LOC340204-specific quantitative RT-PCR (primer pair SEQ ID NO: 107 and 108). LOC340204 is selectively expressed in healthy tissues and strongly overexpressed in some tumors. Distinct overexpression in tumor samples in comparison with various normal tissues was detected in particular in gastric, pancreatic, ovarian, lung and esophageal carcinomas (fig. 44).

The extracellular LOC340204 domains may be utilized according to the invention as target structure of monoclonal antibodies.

Example 29: Identification of LOC340067 as a diagnostic and therapeutic cancer target

The gene which is localized in the LOC340067 gene locus on chromosome 5 (5q22.3) and has the transcript of SEQ ID NO: 109 encodes the protein with SEQ ID NO: 110. The transmembrane protein displays no homologies to other protein domains.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing a quantitative RT-PCR (primer pair SEQ ID NO: 111 and 112) specific for LOC340067. LOC340067 is selectively expressed in healthy tissues and strongly overexpressed in some tumors (fig. 45). Distinct overexpression in tumor samples in comparison with various healthy tissues was detected in particular in pancreatic, mammary, liver, ovarian, lung and kidney carcinomas.

The extracellular LOC340067 domain may be utilized according to the invention as target structure of monoclonal antibodies.

Example 30: Identification of LOC342780 as a diagnostic and therapeutic cancer target

The gene which is localized in the LOC342780 gene locus on chromosome 18 (18q21.32) and has the transcript of SEQ ID NO: 309 encodes the protein with SEQ ID NO: 310. The transmembrane protein includes an acyltransferase domain which is present in many C. elegans proteins which have previously not been characterized in detail.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples (pool of samples, the number is indicated in the figure) was investigated after establishing an LOC342780-specific quantitative RT-PCR (primer pair SEQ ID NO: 311 and 312). LOC342780 is very selectively expressed in healthy tissue, with specific transcripts being detectable especially in the prostate, stomach, testis, lung and the mammary gland (fig. 46). In contrast, LOC342780-specific expression was

surprisingly detected in all types of tumors analyzed. Tumor-specific overexpression of LOC342780 is seen in particular in mammary, ovarian, kidney and liver carcinomas (fig. 46).
LOC342780 is a selectively expressed antigen which is obviously increasingly expressed in proliferating tissues. Thus selective overexpression in tumors can be observed which is therapeutically utilizable. The extracellularly located amino acids 76-89, 316-345, 399-493 and 650-665 (based on SEQ ID NO: 310) may be utilized according to the invention as target structures of monoclonal antibodies.

Example 31: Identification of LOC339511 as a diagnostic and therapeutic cancer target

The sequence according to SEQ ID NO: 113 is derived from a gene which is located on chromosome 1 (1q23.1). The gene encodes the protein of SEQ ID NO: 114. The transmembrane protein displays homologies to the group of olfactory 7-transmembrane receptors.
According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing a quantitative RT-PCR (primer pair SEQ ID NO: 115 and 116) specific for LOC339511. In healthy tissues, LOC339511 is selectively expressed in the liver (fig. 47A). In the carcinoma samples, LOC339511-specific transcripts were identified in liver tumors, with weak expression being moreover detectable in colon carcinomas, mammary and lung carcinomas. When comparing liver-specific expression in tumor and in healthy tissue, increased expression was detected in some tumor samples (fig. 47B).
The extracellular domains of SEQ ID NO: 113 may be utilized according to the invention as target structures of monoclonal antibodies. In particular, the extracellularly located amino acid residues 1-23, 82-100, 167-175 and 226-236 are therefore particularly suitable for producing monoclonal antibodies.

Example 32: Identification of C14orf37 as a diagnostic and therapeutic cancer target

C14orf37 (SEQ ID NO: 125) is a gene which is localized on chromosome 14 (14q22.3) and which encodes the gene product with SEQ ID NO: 126. The transmembrane protein displays no homologies to previously known proteins.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing a quantitative RT-PCR (primer pair SEQ ID NO: 127 and 128) specific for C14orf37. C14orf37 is expressed in various healthy tissues, and strongest in testis (fig. 48). A distinct overexpression in comparison with various healthy tissues was detected in particular in kidney carcinomas.

The extracellular domain of SEQ ID NO: 126 may be utilized according to the invention as target structure of monoclonal antibodies.

Example 33: Identification of ATP1A4 as a diagnostic and therapeutic cancer target

The ATP1A4 gene (SEQ ID NO: 129) is located on chromosome 1 (1q21-23). The gene codes for a protein with SEQ ID NO: 130. ATP1A4 is an integral transmembrane protein with eight transmembrane domains, which is located in the plasma membrane. ATP1A4 is part of a protein complex, with the catalytical part of the sodium/potassium ATPase being present at the N terminus (Woo et al., J. 2000. Biol. Chem. 275, 20693-99). ATP1A4 displays strong homologies to numerous other representatives of the cation ATPase family.

According to the invention, the amount of gene-specific transcripts in healthy tissue and in carcinoma samples was investigated after establishing an ATP1A4-specific quantitative RT-PCR (primer pair SEQ ID NO: 131 and 132). In healthy tissues, ATP1A4 is selectively expressed especially in testis (fig. 49). Strong over-expression of ATP1A4 was detected in some tumor samples in comparison with the respective healthy tissue.

Distinct overexpression in tumor samples in comparison with healthy tissues was detected in particular in pancreatic, mammary, liver and kidney carcinomas (fig. 49), with expression in pancreatic and mammary carcinomas being very high over all.
The extracellular domains of ATP1A4 may be utilized according to the invention as target structure of monoclonal antibodies. The following amino acid residues, based on SEQ ID NO: 130, are located extracellularly: amino acid residues 129-137, 321-329, 816-857 and 977-990.

Example 34: Identification of SEQ ID NO: 133 to 264 as a diagnostic and therapeutic cancer target

The sequences according to SEQ ID NO: 133-266 are 33 genes (nucleic acid sequence, amino acid sequence), together with the respective PCR primers for specific RT-PCR reactions. All proteins have one or more transmembrane domains, but there is little information on homologies to protein domains.
According to the invention, the amount of the particular gene-specific transcripts in healthy tissue and in carcinoma samples was investigated for these genes in specific quantitative RT-PCR reactions. For all of the genes, overexpression which was partially strong in comparison with the respective healthy tissue was detected in tumor samples.
All genes of this group are therapeutically and diagnostically utilizable. The extracellular domains may be utilized here according to the invention as target structure of antibodies.